



PHD

**A Content Based Approach for Investigating the Role and Use of Email in Engineering Design Projects**

Wasiak, James

*Award date:*  
2010

*Awarding institution:*  
University of Bath

[Link to publication](#)

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

**Take down policy**

If you consider content within Bath's Research Portal to be in breach of UK law, please contact: [openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk) with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

**A Content Based Approach for Investigating the Role and Use of  
Email in Engineering Design Projects**

James Oliver Wasiak

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Mechanical Engineering

June 2010

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

A handwritten signature in black ink, appearing to be 'J. Wasiak', with a long horizontal stroke extending to the right.

James Wasiak

**Copyright**

Attention is drawn to the fact that copyright of this thesis rests with its author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and they must not copy it or use material from it except as permitted by law or with the consent of the author.

**Declaration**

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Bath. The work is original except where indicated by special reference in the text and no part of the thesis has been submitted for any other degree.

## **Abstract**

The use of email as a communication and information sharing medium in large, complex, globally distributed engineering projects is widespread; yet there exists little understanding of the content of the emails exchanged and the implications of this content on the design project, design records and contracts. The importance of these issues is underlined by the fact that email records can now be required as evidence in legal disputes. It follows that the overall aim of this research is to assess the role and use of email in engineering design projects.

A state-of-the-art review of literature pertaining to email is reported, along with a review of information and communication processes in engineering design projects.

The primary contribution of this thesis is the creation of a content based approach for analysing the role and use of email in engineering design projects. This centres on the development and application of a coding scheme to email text, identifying what subject matter an email relates to, why it was sent, and how its content is expressed. Results are then analysed with respect to the frequencies of each code and other variables, including how coding varies between different senders and throughout the project duration.

The second key contribution of this thesis is the analysis of emails and content in an engineering setting by applying the aforementioned approach to two case studies. The major case study concerned a large, complex, globally distributed, multimillion pound systems engineering project, from which 16 000 emails were obtained.

It was found that emails are mainly used to transfer information but also to support management functions. Emails facilitate design work but little of this takes place explicitly in the email content. Characteristics of a project affect the subject matter of emails but have little effect on why

they are sent. User roles and personal preferences also influence email use.

If was found that the purposes for sending emails varied over the duration of a project; it was further determined that these changes could be used to identify project progress and design activity.

Implications of the findings are identified in relation to: information management, knowledge management, project management, collaboration and email practice. Significantly, emails do contain potentially important design information and because these often support decisions made elsewhere, emails should be integrated with wider records. More consideration and training should be given to the use of project standards for email use and guidelines for composition. Changes in email use over the project duration could be a potential tool for project managers to identify design progress and possible issues in a project.

## **Acknowledgements**

This work was funded by the KIM Grand Challenge Project and IdMRC at the University of Bath, both of which are supported by the Engineering and Physical Sciences Research Council.

I would like to acknowledge the enormous amount of the time and effort contributed by my academic supervisors, in particular Dr Ben Hicks. I would also like to thank Dr Andy Dong for the inspiration he provided me with for interpreting my data; not to mention welcoming me in Sydney. I also owe a great deal to Dr Linda Newnes, without whose persuasive skills I would not have entered the world of academia.

I am also grateful to Converteam UK Ltd. and particularly Dr Laurie Burrow, for the access to such extensive email collections and hosting me on many occasions in Rugby in order to analyse them. I also owe particular thanks to Craig Loftus for his technical expertise used to convert huge numbers of coded emails from a foreign format into neatly arranged results in Excel. I am grateful for his patience with my changes of decision, along with his coding of many, many emails. I would also like to thank Hamish McAlpine for his contribution in discussing and coding emails for the development of the scheme.

I would also like to acknowledge my friends and family for their support and my colleagues at the IdMRC whose company I enjoyed enormously. Finally my thanks to my fiancé Claire Bailey, for all your support and encouragement which has seen me complete the arduous task of writing.

# **Content**

Content .....	v
List of Figures .....	viii
List of Tables.....	xii
Abbreviations .....	xiii
1.0 Introduction .....	1
1.1. A Global Communication Revolution .....	1
1.2. Engineering Design .....	2
1.3. Communication in the Engineering Design Process .....	5
1.4. Information Management.....	6
1.5. Knowledge Management.....	9
1.6. Electronic Mail .....	10
1.7. Discussion of Research Requirement.....	16
1.8. Research Questions .....	18
1.9. Thesis Outline.....	19
2.0 Engineering Design Projects .....	22
2.1. Philosophy and Nature of Engineering Design .....	22
2.2. Design Process and Product Lifecycle .....	29
2.3. Managing Engineering Projects .....	40
2.4. Design Project Teams .....	44
2.5. Communication in Design Teams .....	48
2.6. Information Management.....	52
2.7. Summary .....	62
3.0 Email: State-of-the-Art.....	64
3.1. The working process of dealing with email .....	66
3.2. Secondary use of email generated data .....	81
3.3. Conclusions of Email Literature Survey .....	87
4.0 An Approach for Analysing Emails .....	93
4.1. Review of Existing Research Methods .....	93
4.2. Research Approach .....	98
4.3. Chapter Summary .....	105
5.0 Coding Scheme Development.....	107
5.1. Principles for Scheme Production, Application and Validation	107

5.2.	Method For Scheme Development .....	111
5.3.	Detailing of Scheme Development .....	115
5.4.	Evaluation of the Scheme .....	125
5.5.	Final Scheme.....	128
5.6.	Limitations of Coding Approach .....	135
5.7.	Chapter Summary .....	139
6.0	Pilot Case Study: Approach, Validation and Preliminary Findings ... .....	140
6.1.	Project Context .....	140
6.2.	Application .....	141
6.3.	Discussion of Validation.....	142
6.4.	Results.....	146
6.5.	Chapter Summary .....	157
7.0	Major Case Study: A large engineering project.....	158
7.1.	Project Context .....	158
7.2.	Application .....	162
7.3.	Validation of Coding.....	164
7.4.	Results.....	167
7.5.	Chapter Summary .....	192
8.0	Findings: Discussion and Analysis of Case Studies .....	194
8.1.	Comparison of Case Studies .....	194
8.2.	Characteristics of Emails .....	195
8.3.	Influences of Person, Location, Role and Company on Email Use .....	202
8.4.	Changes in Email Topics with Time.....	210
8.5.	Signatures of Design Activity and Project Progress.....	212
8.6.	Summary .....	218
9.0	Discussion.....	221
9.1.	Information Management.....	223
9.2.	Knowledge Management.....	227
9.3.	Project Management.....	229
9.4.	Collaboration .....	233
9.5.	Email Practice.....	237
9.6.	Summary .....	239

10.0	Conclusions.....	242
10.1.	Contribution to Knowledge .....	243
10.2.	Key Findings .....	244
10.3.	Research Question 1: What is the nature of the information engineering teams communicate using email?.....	245
10.4.	Research Question 2: How does email support project management and design activity? .....	247
10.5.	Research Question 3: How can an understanding of the content be used to support engineering design projects? .....	250
10.6.	Project Objectives .....	252
10.7.	Further Work and Research Opportunities.....	255
	References .....	i
	Appendix A Semi-Structured Interview Model.....	xxii
	Appendix B – Five Step Guide To Email Coding.....	xxvii
	Appendix C – Author’s Related Publications.....	xxxviii



## **List of Figures**

Figure 1-1: The growth of email use since its conception .....	11
Figure 2-1: A solution generation process (Pahl and Beitz, 1996) .....	25
Figure 2-2: An IBIS style argumentation map produced in Compendium (Buckingham Shum, 2009) .....	26
Figure 2-3: The design process (Pahl and Beitz, 1996 pg. 56) .....	30
Figure 2-4: The Vee Process of System Design and Integration. Based on Forsberg and Mooz (1992). .....	33
Figure 2-5: The total design process model (Pugh, 1991) .....	35
Figure 2-6: Commitment of manufacturing cost during design (Ullman, 2002).....	38
Figure 2-7: Commitment of costs and occurrence of costs in the CADMID Lifecycle (Ministry of Defence, 2009) .....	38
Figure 2-8: Concurrent and sequential engineering processes (Backhouse, et al., 1996) .....	45
Figure 2-9: Project Style Team Structure.....	47
Figure 2-10: Matrix Style Team Structure .....	47
Figure 2-11: Spectrum of asynchronous and synchronous communication methods .....	49
Figure 2-12: An information lifecycle, based on SIMS (2003) .....	56
Figure 3-1: The relevance of various disciplines to topics of email research.....	65
Figure 3-2: A timeline summarising some of the investigations into retrieval and rule based methods for email classification.....	74
Figure 3-3: A screen shot of a 'Themail' word exchange showing one column of words per month (Viegas, et al., 2006).....	83
Figure 3-4: The two sections of the review and the objectives now discussed in conclusion. ....	87
Figure 4-1: Research approach for the study of email use in engineering design projects.....	99
Figure 5-1: Process of coding scheme development, validation and finalisation. ....	112

Figure 5-2: Bales categories for analysing interaction processes, adapted from Bales (1950) .....	123
Figure 5-3: Coding What topics the email discusses: the general theme is discussion of the product, and terms highlighted within this category are components and functions. ....	130
Figure 5-4: The coding of why emails are sent: the purposes of confirming and managing are highlighted.....	132
Figure 5-5: The coding of how email content is expressed using terminology from Bales. ....	134
Figure 5-6: A threaded email conversation viewed in Google™ Mail....	136
Figure 5-7: A quoted text email reply shown in the University of Bath's webmail client. ....	136
Figure 6-1: Distribution of email lengths in the Pilot Case Study corpus .....	141
Figure 6-2: The proportions of emails relating to the three topic groups and combinations thereof.....	147
Figure 6-3: The percentage of emails across the sample coded by terms from within the Product group. ....	148
Figure 6-4: The percentage of emails across the sample coded by terms from within the Project group. ....	149
Figure 6-5: The percentage of emails across the sample coded by terms from within the Company group. ....	149
Figure 6-6: The proportions of emails relating to the three purposes for sending emails and combinations thereof.....	150
Figure 6-7: The percentage of emails across the sample coded by the different elements which comprise the three purposes for sending emails. ....	151
Figure 6-8: The percentage of emails from across the sample coded by terms which describe how content is expressed.....	152
Figure 6-9: The proportion of Product, Project and Company coded emails sent by four key users in the project. ....	153
Figure 6-10: The proportion of Information, Management and Problem Solving Purposes coded in emails sent by four key users in the project. ....	154

Figure 6-11: The change in the frequency of emails relating to the three topic groups over the project duration.....	154
Figure 6-12: The change in the frequency of emails relating to the three email sending purposes over the project duration.....	155
Figure 7-1: Parties involved in the project.....	159
Figure 7-2: The organisation of the project and roles within the company. ....	160
Figure 7-3: Distribution of email lengths in the Major Case Study corpus .....	162
Figure 7-4: The proportions of emails relating to the three topic groups and combinations thereof.....	168
Figure 7-5: The percentage of emails across the sample coded by the terms within the Product group. ....	169
Figure 7-6: The percentage of emails across the sample coded by terms from within the Project Group.....	170
Figure 7-7: The percentage of emails across the sample coded by the terms from within the Company group. ....	170
Figure 7-8: The proportions of emails relating to the three purposes for sending emails and combinations thereof.....	171
Figure 7-9: The percentage of emails across the sample coded by the different elements which comprise the three purposes for sending emails. ....	172
Figure 7-10: The percentage of emails from across the sample coded by terms which describe how content is expressed.....	174
Figure 7-11: The proportion of Product and Project coded emails sent by key users in the industrial partner company.....	175
Figure 7-12: The proportion of Product and Project coded emails sent by key users from major parties in the project. ....	176
Figure 7-13: The proportion of Information, Management and Problem Solving purposes coded in emails sent by key users in the industrial partner company.....	177
Figure 7-14: The proportion of Information Transactions, Management Transactions and Problem Solving Contributions coded in emails sent by key users from major parties in the project. ....	179

Figure 7-15: The change in the frequency of emails relating to the three topic groups over the project duration.....	181
Figure 7-16: The change in frequency of emails coded to terms within the Product group over the project duration.....	182
Figure 7-17: The change in frequency of emails coded to terms within the Project group over the project duration, part 1.....	184
Figure 7-18: The change in frequency of emails coded within the Project group over the project duration, part 2.....	185
Figure 7-19: The change in the frequency of emails relating to the three purposes for sending emails over the project duration.....	187
Figure 7-20: The change of frequency of emails coded within the Information Transaction Group over the project duration.....	188
Figure 7-21: The change in frequency of emails coded within the Problem Solving Contributions Group over the project life.....	189
Figure 7-22: The change in frequency of emails coded within the Management Transaction Group over the project duration.....	190
Figure 8-1: The proportion of Product and Project coded emails sent by key users in the industrial partner company. (Reprint of Figure 7-11)...	202
Figure 8-2: The proportion of Product and Project coded emails sent by key users from major parties in the project (Reprint of Figure 7-12). ....	203
Figure 8-3: The proportion of Information Transactions, Management Transactions and Problem Solving Contributions coded in emails sent by key users from major parties in the project (Reprint of Figure 7-14). ....	204
Figure 8-4: The proportion of Information, Management and Problem Solving purposes coded in emails sent by key users in the industrial partner company (Reprint of Figure 7-13).....	205
Figure 8-5: Change in the frequency of emails relating to the three purposes for sending emails over the project duration (Reprint of Figure 7-19). ....	213
Figure 8-6: The change in the frequency of emails relating to the three topic groups over the project duration (Reprint of Figure 7-15).....	214

## **List of Tables**

Table 4-1: A summary of relevant methods for the analysis of email use discussed in section 4.1. ....	97
Table 5-1: Initial coding scheme for What topics were discussed. ....	116
Table 5-2: Development of coding scheme structure for what topics were discussed. ....	118
Table 5-3: Terminology identifying purposes for which communications occur. ....	119
Table 5-4: Three proposed groupings of communicative purposes .....	121
Table 5-5: Coding terminology and definitions identifying what topics email discuss. ....	129
Table 5-6: Coding terminology and definitions identifying why emails are sent, with respect to transactions or contribution. ....	131
Table 5-7: Coding terminology and definitions used to identify characteristics of how email dialogue is expressed. ....	133
Table 6-1: Inter-coder reliability of top level codes calculated using Cohen's kappa. ....	143
Table 7-1: Roles descriptions for the project team members interviewed. ....	164
Table 7-2: Inter-coder reliability of top level codes calculated using Cohen's kappa. ....	165
Table 7-3: Statistical Significance of Top Level Codes Using the Spilt Half Technique .....	166
Table 7-4: The stability of top level codes checked five days apart .....	166
Table 7-5: The percentage of emails for each combination of purpose and topic. ....	173
Table 8-1: A Comparison of Characteristics of the Projects from the Pilot and Major Case Studies. ....	195
Table 8-2: Factors affecting the use of email .....	210
Table 8-3: Permutations of rises and falls in information, management and problem solving purposes and interpretations, drawn from Figure 8-5. ....	215

Table 9-1: A summary of key suggestions and observations to improve email use discussed in this Chapter.....	222
Table 10-1: A summary of the key findings from the two case studies. ....	244
Table 10-2 Possible further terms identified during coding .....	257

## **Abbreviations**

CAD – Computer Aided Design / Drawing  
 IBIS – Issue Based Information Systems  
 IM – Information Management  
 IR- Information Retrieval  
 IT – Information Technology  
 KM – Knowledge Management  
 NB – Naïve Bayes  
 NLP – Natural Language Processing  
 PDM – Product Data Management  
 PLM – Product Lifecycle Management  
 PM – Project Management  
 RF – Random Forest  
 SVM – Support Vector Machines  
 TFIDF – Term Frequency Inverse Document Frequency

## **1.0 Introduction**

The overall aim of this research is to investigate the role and use of email in engineering design projects through an analysis of email content. The use of email as a communication and information sharing medium in large, complex, globally distributed engineering projects is widespread; yet implications of its use in preference of other communication media and its integration with other project records has not been properly assessed. The need to assess the role and use of email is recognised by the industrial partner for this research, Converteam UK Ltd. To support this argument a discussion of communication, information and knowledge management in the context of engineering design projects is now presented. Attributes of email are explored followed by a discussion of the research requirement. The research questions and objectives are then posed.

### **1.1. A Global Communication Revolution**

A global communication revolution has changed the way economies function and made close working relationships between individuals at other ends of the earth not only possible but practical and a social normality (Cairncross, 1997). This paradigm has been exploited by the creative industries (Caves, 2000), for whom the ability to share and distil information as effectively as possible is central to their *modus operandi* (Castells, 2000; Flew, 2002). These industries (including product design, arts, media and publishing) require knowledge working, that is to say individuals generating ideas and developing concepts, before being crafted by more traditional manual work processes into tangible artefacts (Hartley, 2005). The asset of intellectual property is vital to these companies along with the ability to safely share and generate it. Communication and information sharing is therefore central to knowledge work, with distance becoming less relevant as technologies such as email

are used equally for communicating with the person next door and with colleagues and customers in other countries.

Innovative firms acknowledge the benefits of global collaboration for successful product creation (Chesbrough, 2006) particularly in the case of engineering design. In this domain, where products and technical systems are conceived, challenges are at their most critical often involving hundreds of engineers as knowledge workers, based in numerous locations and countries, operating to multimillion pound budgets on safety critical systems such as aircrafts, automobiles and machine systems (Norris, 2005).

## **1.2. Engineering Design**

Engineering design can be understood to be a process of creatively applying scientific principles combined with technical understanding and expertise to realise the design of an artefact or system (Pahl and Beitz, 1996; Dorst and Cross, 2001). It is often an iterative and open ended process and may need to consider numerous factors relating to the lifecycle of the product, such as sourcing of materials, manufacturing, transportation, disposal or decommissioning and of course the “in use” part of its life. Engineering design can therefore encompass a broad spectrum of expertise requirements and range from relatively simple and quick, through to long life and complex designs. Throughout the diverse range of processes enacted during engineering design, effective communication between members of the design team is becoming ever more important and pervasive (Badke-Schaub and Frankenberger, 1999). Reasons for this include increasing product complexity, reduced lead times to product launches, the adoption of concurrent engineering methodologies, and a shift towards globalisation and product service systems. These are now discussed.



## **Complex Products**

Historically the design process has been conducted by a few individuals working within their domain of expertise, to design a product of which they would have full understanding; consequently there was very little need for documenting or sharing of information beyond the organisation or department (Clarkson and Eckert, 2005). This is very different from the world today, where increasingly complex products such as aircraft will have hundreds of thousands of parts, be designed by tens of thousands of engineers from a wide range of disciplines, with a requirement to communicate large volumes of design information between themselves (Faulconbridge and Ryan, 2002; Boeing, 2009).

## **Product Lifecycle**

Designing is typically understood to be the process of moving from a problem statement through to the output of a detailed design, via stages of concepts, developments and refinements (Zeid, 1991; Pahl and Beitz, 1996). Following this, the stages of manufacture, sales, distribution, service and disposal will typically occur. Increasingly parties involved in the various stages of this lifecycle will need to share information with one another (Backhouse, et al., 1996; Saaksvuori and Immonen, 2008), and hence communication and collaboration between the design team is vital (Kleinsmann and Valkenburg, 2008).

## **Concurrent Engineering**

With customer demands requiring higher product quality at lower cost, concurrent engineering aims to maximise profit through reducing the time to market of new products (Prasad, 1995). This is achieved by the overlapping stages of the development process which would traditionally have run sequentially (Backhouse, et al., 1996). Whereas previously an activity would have been completed before being passed on for the next activity to start (commonly referred to as the over the wall method (Ullman, 1992)) designers are now required to communicate with colleagues already working on the following activity advising them of

decisions and changes. This method places increased importance upon effective and accurate communication and information transfer, much of which may be undertaken by electronic means.

### **Globalisation**

Globally distributed working has become common in engineering where there is a trend for design activities to be conducted in the UK with manufacturing taking place elsewhere (e.g. Triumph Motorcycles manufacture in Thailand (British Chamber of Commerce Thailand, 2009)) or designed in many places, built in many places, assembled in one place (e.g. Airbus and the A380 (Airbus, 2009)). In particular global companies can distribute work to take advantage of different local economic factors, such as lower labour costs for manufacturing. Building of local customer relationships and understanding cultures may also offer competitive advantage (Porter, 1990). Although globally distributed working has been largely driven by these economic factors, it is advances in electronic communication and information technology (IT) which have enabled this (Prusak, 2001; Friedman, 2005).

In order to enable globally distributed working two challenges have needed to be overcome; firstly communicating over long distances and secondly across time zones. The potential for electronic communication technology to provide fast, data rich yet asynchronous transmission was recognised as early as 1995 (Fulk and DeSanctis), before email via The Internet was in mainstream use. Email is now one of, if not the dominant communication channels of business worldwide (Wilson, 2002; The Radici Group Inc., 2008).

### **Product Service Paradigm**

There is an emerging paradigm shift in business strategy by engineering firms (Davies, et al., 2003; Oliva and Kallenberg, 2003), which is currently the subject of much investigation (Ball, et al., 2006). Traditionally companies sold products to a customer who would then be responsible

for their own maintenance arrangements and associated costs. In the emerging product service market, engineering companies sell the use of a system to the client and support the product through its service life and later disposal (Castells, 2010). This has the advantage for the customer of guaranteeing performance of a system for a known cost (Baird, et al., 2000). For the engineering organisation there is opportunity to realise profit beyond the sale of the original product. In the long term this paradigm should encourage engineers to produce better products as there is a self-motivated interest in their successful operation and an opportunity to learn through life how to improve future designs. Critical to this long term approach is the management of information and knowledge to meet the needs of the future (Sivaloganathan and Shahin, 1998; Tallon and Scannell, 2007), whatever they may be, and to feedback information from in-service products to design teams (Goh and McMahon, 2009).

The emergent trends of increasingly complex products, reduced time to market, concurrent engineering, globalisation and product service models put great emphasis on the management of communication, information, and knowledge sharing. These dimensions are now discussed with particular reference to engineering design.

### **1.3. Communication in the Engineering Design Process**

The Oxford English Dictionary (2005) defines communication as ***“the imparting or exchanging of information by speaking, writing, or using some other medium.”*** For the purpose of this work it is useful also to define the term “Information Sharing” which describes the act in communication of pointing or directing to an information source to enable its location and retrieval. In general communication the sender actively conveys the information of interest whereas in information sharing the information resource to be shared plays a passive role in the communication process. This is more constrained than definitions generally used (e.g. Malone, et al., 1987; Constant, et al., 1994) but

necessarily distinct for this work. Information sharing can also include associated acts, such as giving digital permissions for the intended recipient to access the resource.

The relationship between communication and information is important for the following discussion as these terms are central in literature describing how designers operate (e.g. Dong, 2005; Törlind, et al., 2005). The collaborations between parties involved in design necessitate the effective communication and sharing of information. Hence design is considered to be a ***“knowledge intensive process of communication”*** (Darlington, 2002) with players requiring information and understanding of scientific principles as well as about the problem (Pahl and Beitz, 1996). The effective, timely and accurate communication of this knowledge and information is a fundamental part of the process.

***“Communication is an essential part of any design process”***  
***(Eckert, et al., 2005)***

### **1.3.1. Project Management and Design Teams**

In order to support the communication requirements of engineering designers, involved in increasingly complex, distributed, multidisciplinary and long life projects; a wave of technologies and techniques has emerged under the guise of “Information Management (IM) and Knowledge Management (KM).” These are now discussed.

## **1.4. Information Management**

The terms Information and Knowledge Management have taken various meanings by different authors (Hicks, et al., 2002; McMahon, et al., 2004), even within the domain in which they have worked (Huet, 2006). These varied ideas, explored in more detail in Chapter 2, are defined here.

**Information Management** (IM) encompasses such activities as organising, accessing and sharing information. Standards such as the format and layout of documentation prescribed by an organisation are also encompassed. The objective of IM is to make most effective use of information resources.

The concept of **Knowledge Management** (KM) is more ambiguous being that knowledge is not a quantifiable or physical asset. KM is considered here to involve managing the environment in which knowledge is created generated and shared. It is also recognised that because information is a prerequisite for knowledge processes IM will affect KM.

IM has become heavily entwined with IT and the development of Information Systems to support their specific user groups (Galliers, et al., 1999). One of the main developments in engineering has been the implementation of Product Data Management (PDM) Systems (Kahn, 2005), which help control, maintain and support the access of, and relationships between different items of product data, such as Computer Aided Design (CAD) models, analyses, specifications, reports and manufacturing information. Although the information systems side is well developed, management challenges still lie in the effective retrieval of information, which has been the subject of much investigation, e.g. (Hicks, et al., 2002; Li and Ramani, 2007). Further difficulties include determining the quality or value of information (Zhao, et al., 2007) and the stage in development or maturity of information (Grebici, et al., 2007). It is also recognised that an increasing problem is the “overload of information” incurred by those searching and their difficulties in identifying ‘what is’ and ‘is not’ of relevance (Eppler and Mengis, 2004). Structured or formal information sources, such as CAD models or document databases, are typically seen to be more helpful in this regard as their organisation or classification allows browsing, rather than relying on keyword searching alone. Less formal and less structured supporting information artefacts, such as lone paragraphs of text on Post-it® notes or sketches without context, are more difficult to re-use and retrieve

(Gunendran, 2007). This is primarily because the technologies used to manage formal information cannot be so readily applied to less structured information.

These less formal sources of information, of which email is included, are increasingly recognised for their value. Lowe et al., (1999) identify that activities undertaken in the early stages of design are often more heavily reliant upon the less structured or informal information types. It is recognised that the information needs of the engineer will evolve as a project progresses and moves through its service life, making it difficult to predict the future relevance of information sources (Eckert, et al., 2007). Indeed at any point in time engineers need to access information for a wide variety of reasons, such as problem solving, understanding situations, determining product features and presentation to decision makers (Kwasitsu, 2003). The reasoning for making decisions, referred to as rationale, is recognised as important to capture. Shipman and McCall (1997) recognise the value in recording communication which contains rationale as no extraneous processes need to be undertaken by engineers, which would risk affecting their working behaviour. Contrary to this, the preparation of documentation requires engineers to record their activities in a way which is most useful to the intended recipient, rather than capturing both the language and form of what actually happened. An alternative technique is to use a tool such as DRed (Bracewell and Wallace, 2003; Bracewell, et al., 2009) with which designers create an annotated diagram, mapping the argumentation of a design scenario.

It is arguable that a large amount of rationale is exchanged during communication processes such as meetings, telephone calls, video conferencing and email (Huet, 2006). Furthermore, face-to-face meetings and telephone conversations are seen to be very “information rich”, because of their intimate feedback and use of tone of voice. This concept of Information Richness (Daft and Lengel, 1986) is defined as “the ability of information to change understanding within a time interval” interpreted as the effectiveness and depth of media in communication. Daft and

Lengel (1986) determine that impersonal documents, and numeric representations typical of engineering design are of low richness, whereas personal letters which include email, are seen to perform more strongly. Further to this, the ability of email to support rich media was recognised early in its emergence (Lee, 1994) and more recently (Panteli, 2002). Hence, there is strong evidence to suggest that electronic communication and in particular, email may include a large amount of potentially valuable information including design rationale.

## **1.5. Knowledge Management**

It has been shown that it is often the case that people rely on their knowledge, and the knowledge of those around them in order to solve problems (Heiman, et al., 2009). Consequently, these dependencies can cause difficulties when people leave an organisation. By identifying and making visible how communities share knowledge it is proposed that these difficulties can be reduced (Cross, et al., 2001). It has been recognised that patterns of email use have a potential role in identifying and locating expertise and knowledge sharing relationships within organisations. Tyler et al., (2005) use email logs to find communities of practice, which Campbell (2003) suggests is sufficient to locate expertise without considering the content of the emails. Individuals have been shown to use documentation to identify people who may be able to help them, and also to use people to find appropriate documentation (Hertzum and Pejtersen, 2000). This emphasises the potential value in exploiting community structure information from email, such as who speaks most frequently with whom.

In addition to considering the structures of communities the language used between individuals or teams in their communication can provide valuable insights into designers' cognitive processes i.e. the way in which they develop knowledge and arrive at solutions (Dong, 2005). The application of automatic tools to analyse meeting discourse, such as

Agoraprobe (Dong and Vande Moere, 2006) have been shown to identify the process of knowledge creation and level of collaboration within teams. In a similar way email communications have been shown to support knowledge development and creativity in a purposeful way (Lichtenstein, 2004).

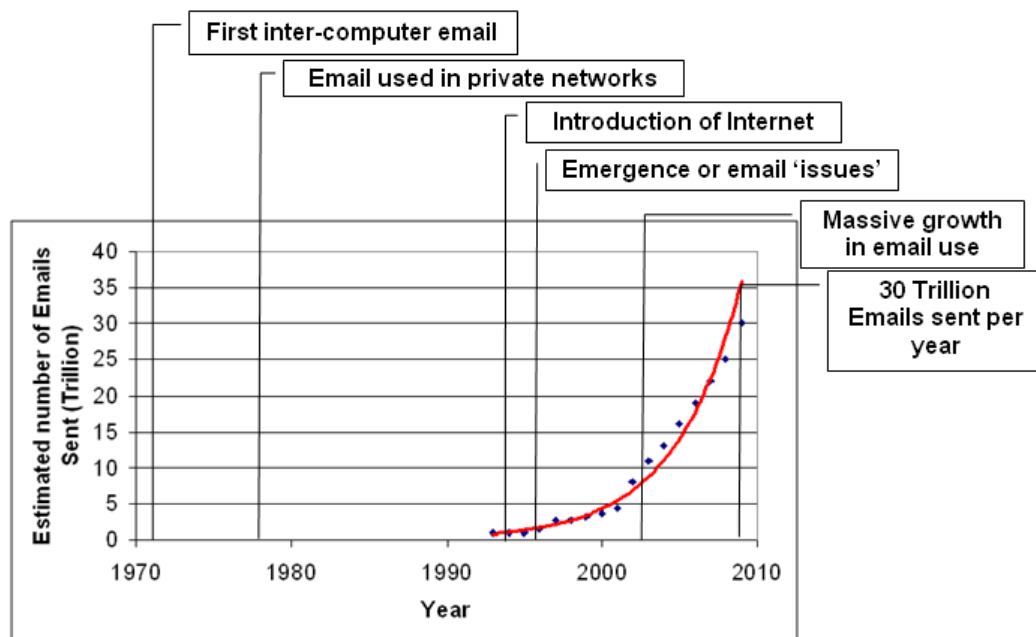
## **1.6. Electronic Mail**

The previous sections have discussed the activities of Information and Knowledge Management and highlighted the potential uses of email as a part of holistic IM and KM strategies. The significance of email in the context of engineering design has also been alluded to. Despite this relatively little research has been conducted in the domain of email; and none specific to engineering design. Given the potential benefits and lack of existing research there is arguably a strong case for exploration of this communication media. An overview of email is now presented, with further reference to its relevance in engineering design and a summary of current understanding.

Email or Electronic mail has grown to become a dominant communication medium for both business and private individuals in recent years. Over 10 trillion person to person emails and a further 15 trillion 'alert' and SPAM messages were predicted to be sent in 2008 (Gantz, et al., 2007). Figure 1-1 illustrates this email growth based on a range of available data.

The origins of email can be traced back to the late 1960's and early 1970's but it was to be sometime before its potential was realised and its use spread, requiring the development of the Internet (Leiner, et al., 1997). Initially its use was limited to computer development and academic communities, but as connection to the Internet spread, spurred by the introduction of the World Wide Web in the mid 1990's, email began to take off (Turner and Ross, 2000).





**Figure 1-1: The growth of email use since its conception**

Email has been considered a ubiquitous office tool since the turn of the century (McManus, et al., 2002), with increasing use propelled by technologies such as the Blackberry mobile phone (Mazmanian, et al., 2006). Email is involved in core processes in many organisations. Processes such as invoicing, handling customer inquiries, discussion of key strategy and contract negotiation have each been shown to be handled by email in over 60% of organisations (AIIM International, 2003). As a result hundreds of thousands of emails may be generated by organisations and project teams, containing potentially unique references to discussions, data, decisions and reasoning which may need to be retrieved at a later date.

One survey suggests that almost half of office based employees spend in excess of 30% of their day involved in email related activity (AIIM International, 2006). It has also been demonstrated that for many users, email has become central in functioning as a task management resource (Bellotti, et al., 2003). Clearly this has far reaching implications for project management (PM) and underlines how central email has become to modern working processes.

Yet despite this wide and varied use, organisation's responses to email management and policy appear unclear. Over a third of email users report their organisations have not begun implementing an email management strategy (AIIM International, 2006). Given the significance of email this is somewhat concerning, as failure to manage this resource correctly may have consequences not only in terms of business but in the case of product design in terms of safety. The following quote from the partner company to this research identifies their main concern ***“Email is seen as a surrogate for private conversation, and this view of email as a private communication medium may inadvertently preclude the pursuit of group information management based on private email for the benefit of project management.”***<sup>1</sup>

#### ***1.6.1. The direct and indirect cost of email***

The cost of communication can be considered in terms of direct costs, such as a telephone bill or postage stamp and indirect costs, such as the time taken to write a message. As such, email use tends to incur only indirect costs associated with sending (or receiving), unlike telephone and fax where direct costs also apply. Broadly speaking emails incur IT system support costs and human resource costs. Because email requirements overlap with other IT provisions, the proportion of the costs which they bear would be very difficult to allot. An exception to this may be the storage space related to email management, of particular relevance to long term record keeping. The “footprint” (backups, temporary data, duplication) associated with an email can be significantly greater than the original email size (Gantz, et al., 2008) and will increase proportionally with the number of emails stored. For example a one megabyte email sent to four people could use around fifty megabytes including all temporary files and backups. It may therefore be worth the

---

<sup>1</sup> Personal communication with Dr L D Burrow, Converteam UK Ltd, 24<sup>th</sup> July 2009.

investment of organisations to determine which emails to keep and which to delete, minimising storage costs (Carver, et al., 2006).

Although emails facilitate communication they also have the affect of reducing human contact and incur human costs due to interruptions, overload and a pressure to always be online (Day, 2009). As such the human cost of email may be as difficult to calculate as the technical cost. It is possible to record how much time employees are working on email related activities but this does not give an indication of how necessary or valuable this effort is. Jackson et al., (2006) attempt to consider this by measuring employees time spent reading email and using a number of averaged factors to consider what proportion of this is wasted time. The utility of this numerate approach is debatable, however, their suggestions that improved targeting of emails, construction of email content and implementation of user training will improve email user efficiency are seen to be valid.

Also related to the question of how much time spent dealing with email is useful, there is the issue of overload; engineers and particularly managers can receive as many as two hundred emails a day<sup>2</sup>, and consequently feel overloaded (Carver, et al., 2006). Furthermore recipients tend to perceive more pressure to act upon an email than senders actually require (Renaud, et al., 2006).

It is therefore argued that there are opportunities to improve efficiency with which engineering teams interact with email. However to achieve this it is necessary to have an understanding of how emails are used in an engineering project setting, something not afforded by current research.

---

<sup>2</sup> Communications between Dr Ben Hicks, University of Bath and Jeremy Watson, Director, Global Research, Arup, 2007.

### **1.6.2. Emails and Record Keeping Strategy**

Regardless of whether it is appropriate and whether it is desired, it is acknowledged that most companies regard email as a tool which is part of their record keeping strategy (Pennock, 2006) incurring legal obligations. A number of pieces of legislation including the Data Protection Act (1998), Freedom of Information Act (2000), Regulation of Investigatory Powers Act (2000), Human Rights Act (1998) and Intellectual Property legislation affect companies obligations to preserve and destroy email records (Pedley, 2006; Pennock, 2007). Over the last decade, a number of cases have highlighted the significance of email correspondence in law, and that corporations may be held to account if they fail to have maintained a record of emails required in court (Genborg, 2005). In 2007, two UK supermarket chains were forced to provide over 11 million emails to their industry watchdog, to aid in an investigation (Finch, 2007). Episodes such as this have led to recognition of the legal significance of emails, supported by the finding that in 30% of large organisations the legal department has more influence on email policy than the IT department (Allim International, 2006).

As a communication medium emails are unique in that they represent conversation which may not otherwise have been recorded (if replacing face-to-face meetings or phone) but can also represent information intended to be stored (if replacing a letter or fax with carbon copy). At present there is no standard practice for differentiating between these purposes, requiring one or more parties to pass judgement as to the email's significance in time. This aspect merely adds to the complexity of managing email, although benefit is seen in capturing richer information than before (discussed previously).

The retrieval of information from email records poses a major challenge. In particular the semi-structured format of the information is difficult to deal with computationally and keyword searching alone does not always achieve sufficient results particularly as it cannot deal with context

(Delphi, 2008). Reasons for this include a lack of common practice, training and standard ways of working. To compound difficulties, engineers have a tendency to use terminology and acronyms specific to projects and domains. Hence, there is an opportunity to create better techniques for searching through email archives to aid in information retrieval (IR) and reuse, by improving structure and metadata within emails or email databases, either manually by the sender or recipient, or automatically by the computer.

In engineering, records change with time, versions of designs are superseded and specifications change. Emails often contain links to, or include as attachments, such records. This is usually because it is easy for the user to do so (Cayzer, 2004), but there can be associated complications because of the dynamic nature of these other sources. It is useful to recognise that different users will have different requirements in this regard; sometimes they will want to know what something looked like at a certain time, what it looks like now and what changes have been made. Although PDM systems begin to address the temporal issue associated with the evolution of documents and CAD models there is a lack of current capability to identify and link these changes in any sources referring to these, such as an email.<sup>3</sup>

### ***1.6.3. The Issues of New Technology***

With new technologies, the benefits they offer generally bring with them a new set of problems. Email may be considered to do just that. The affordances of email have been very important for the emerging communication, information and knowledge needs, and management. Email has become one of the primary communication media for business today, with the volume of emails and time spent in email related activity

---

<sup>3</sup> Personal communication with Nigel Hicks, Contracts Manager, Converteam UK Ltd, 9<sup>th</sup> May 2008.

increasing. The speed of communication, even over long distances and its asynchronous nature have meant that email has assisted working across globally distributed teams, whilst at the same time remaining useful as a communication medium within office buildings.

Most research to date has tackled the issues of SPAM (Veloso and Meira, 2006), task management via inboxes (Gwizdka, 2002; Bellotti, et al., 2003), filtering of incoming email (Aery and Chakravarthy, 2005) and identifying expertise or networks (Campbell, et al., 2003; Dom, et al., 2003). Although this range of issues is fairly comprehensive, the level of detail and applicability specifically to engineering design is seen to be lacking, and in this context, key areas have been overlooked. The nature of email has made it amenable to engineers due to its versatility, and it is recognised that much valuable information may be contained in email records.

The way in which email is used by design engineers as a collaborative communication tool has not been addressed, which is considered to be fundamental, given the distributed team working nature of design. The role which email plays in KM and IM, particularly pertinent to long life, complex and distributed projects, has also yet to be fully explored. There is also potential to improve understanding of how networks of people communicate offering benefits to KM strategy. To exploit, enable and realise all of these, it is vital that the role and content of email within engineering design is first understood.

## **1.7. Discussion of Research Requirement**

To summarise and draw from the previous sections, a discussion of the research requirement is now made. It is widely recognised that electronic working and digital communication, such as email, have enabled knowledge working to more easily take place in globalised and distributed settings and that this is becoming increasingly vital to business

(Department for Culture Media and Sport and Department for Business Innovation and Skills, 2009). Email use presents a wealth of opportunities and a wealth of challenges. There is therefore a requirement to capitalise and address these, in order to maximise the effectiveness of engineering design companies and projects.

A number of difficulties with email use however exist, impeding its effectiveness. This includes the overload of messages experienced by some users, difficulties in managing the large email collections and a limited integration with the electronic working environment. It has also been identified that there are opportunities to exploit email use for identifying expertise, assessing team effectiveness and as a method for capturing design rationale not usually contained in formal records. At present however there is a limited understanding of the role and use of email in engineering design settings and in particular its content.

In conclusion of the arguments presented so far, there are two mutually interrelated challenges regarding email use in engineering design and specifically engineering design projects, namely:

- To improve the way in which email is used as part of comprehensive and holistic information management strategies, particularly with regards to archiving and retrieval of information relating to design details and decisions.
- To improve the efficiency and effectiveness of the working processes of designers and support staff in engineering design projects, recognising that email currently plays a key role in many tasks central to normal modes of working.

Prior to addressing these challenges there is a fundamental need to understand the role and use of emails within engineering design projects. In this work the term *role* is defined as: the purpose or objective of an

email and the genre of subject matter or area which it relates to. This is similar to the use of the term by (O'Kane, et al., 2007).

The following section presents the opportunities discussed and elicits research questions and objectives.

## **1.8. Research Questions**

The overall aim of this research is to investigate the role and use of email in engineering design projects through an analysis of email content. The introduction has identified the significance of email use in engineering projects as part of both information management strategies and working processes. To ascertain whether email can be evaluated in terms of engineering design projects the following research questions are posed:

**Research Question 1:** What is the nature of the information engineering teams communicate using email?

**Research Question 2:** How does email support project management and design activity?

**Research Question 3:** How can an understanding of the content be used to support engineering design projects?

In order to investigate these research questions a number of key objectives are identified:

- 1) To examine the existing research and determine what is known about the role that email plays both in engineering design and other working environments. (Chapters 2 and 3)
- 2) To examine the tools and techniques proposed to analyse email use, including archiving and retrieval as well as general work processes. (Chapter 3)



- 3) To develop and validate an approach for analysing the role and use of email within the context of engineering design projects and the overall design process. (Chapters 4 and 5)
- 4) To apply the proposed approach to recently completed industrial projects and to explore the role that email plays over the project life (Chapters 6 and 7)
- 5) To elicit general observations regarding the role and use of emails in engineering projects (Chapter 8).
- 6) To highlight implications and opportunities to improve the way in which email is used to support engineering design projects (Chapter 9).

## **1.9. Thesis Outline**

**Chapter 2 “Engineering Design Projects”** - begins with an overview describing the nature of design activity and how this is typically supported by engineering organisations. The process of designing and phases throughout product and project lives are then discussed. Key aspects in the management and control of engineering projects are then outlined. The use of teams and effective team working in engineering projects is then reviewed, with reference to distributed working and managing teams. Information management in design projects is discussed and challenges highlighted. A general background to communication in design settings is presented, which leads to a discussion as to the criticality, difficulties and opportunities of email.

**Chapter 3 “Email – State of the Art”** - provides a far reaching review of literature relating to email. The phenomenon of email as a ‘functionally overloaded’ application is discussed, an example of which is its use as a task management system. The resulting issues of having too many emails, and how best to prioritise and file them are also discussed. A detailed review is then presented of research regarding the effectiveness of email for facilitating various functions and processes, especially when

compared to other communication and information sharing mechanisms. Finally, user and research orientated tool developments that aim to make use of the secondary data such as email traffic patterns are presented.

**Chapter 4 “An Approach for Analysing Emails”** – opens with an in depth critique of relevant existing research methods. The first part of this looks at methods which analyse design communication. The second part explores methods for analysing information use in design. These ideas are then brought together and their implications for this study discussed. Based on this an approach is proposed which centres on the analysis of email textual content via coding mark up. The need to collect supporting evidence from interviews and project documentation to use as a secondary data source is also described. An overview of possible analyses is also given.

**Chapter 5 “Coding Scheme Development”** – addresses the rigorous development of a coding scheme to be used in the approach. Theoretical principles for the production, application and validation of a coding scheme, based on existing literature, are presented. The method via which the scheme was then developed is presented; this involves iterative application of coding terminology by a number of researchers to real engineering emails. The details of this development including rationale for key decisions are then reported. The final coding scheme is evaluated with respect to the theoretical principles, to ensure a successful result. The completed coding scheme is presented, along with a guide and examples of how it is to be applied. Finally limitations are discussed.

**Chapter 6 “Pilot Case Study: Approach Validation and Preliminary Findings”** - This chapter outlines a case study which was undertaken for the purpose of validating the overall approach, and also providing some preliminary findings regarding email use in an engineering setting. The context of the project, a small, internal software development project is presented. The application of the approach is then reported and basic

results presented. A review of this application is then given, discussing its suitability. Finally key findings from this study are drawn from the results.

**Chapter 7 “Major Case Study: A large engineering project”** - This chapter details a case study of a large, complex, multinational engineering project. The project background is comprehensively detailed, with particular regard to the chains of communication and project structures both inside and outside of the organisation from whom the corpus was obtained. Details as to how the previously developed approach was applied are given. Results are then presented and discussed only with regards to this specific project. Metrics assessing the significance and validity of findings are discussed here.

**Chapter 8 “Findings: Discussion and Analysis of Case Studies”** - This chapter discusses further the results from the “major engineering project” case study with regards to email use in general engineering settings (i.e. other projects). This includes some preliminary guidance and advice for knowledge and information practitioners.

**Chapter 9 “Discussion”** – This chapter draws on the findings of the case studies and other evidence gathered in the thesis, including the literature review of email. Implications of the findings are identified in relation to: information management, knowledge management, project management, collaboration and email practice.

**Chapter 10 “Conclusions”** - The final chapter returns to the overall aim of this research and highlights the key contributions made. The three research questions are discussed in turn identifying findings from existing literature and case studies, and further implications. Finally further development and research opportunities are discussed.

## **2.0 Engineering Design Projects**

This chapter presents an overview of engineering design projects, providing the relevant context for the investigation and analysis of the role, behaviour and interactions of email in this setting. It also serves to reinforce the arguments for the significance of email in this domain. It begins with an overview of engineering design. Following this design processes and product lifecycles are discussed in more detail. A background of project management is then given followed by an overview of design teams within the context of globalisation. Finally the issues of communication, information and email use in the context of engineering projects are discussed.

### **2.1. Philosophy and Nature of Engineering Design**

The term engineering design has already been used extensively and has been described as an application of creativity, scientific principles and technical knowledge to arrive at a solution. But what distinguishes engineering design from any other type of engineering, or indeed design activity?

From a pragmatic view point, the following example adapted from Ullman (2002) gives a helpful insight. Ullman makes the distinction between engineering as an act of analysis and engineering through design. In an analysis problem, an engineer might have to determine what diameter of a specific bolt type is required to fasten two pieces of metal of specified dimensions and loading, to which there is only one possible and correct solution. In a design problem however, the engineer might be asked to design a joint between two pieces of metal; in this scenario there are many possible methods of making the join (welding, bolting, riveting) and a range of considerations affecting the best choice (operating temperature, assembly location). The choice of join might affect (for example) the minimum thickness of the pieces of metal, which as a

consequence may affect other components in the system. No single correct solution exists in this instance, and several possibilities might be explored to find an optimal result. Hence engineering design is an iterative process, which is usually terminated by time or cost constraints or when an acceptable solution is found.

From a more theoretical perspective, it is argued and indeed widely recognised that engineering design is an archetypal complex problem solving process (Simon, 1969; Goel, 1995). It is characteristic of these problems to be ill-structured and it is hard to determine when and whether they have been solved. The information available to arrive at a solution to these problems is incomplete from the outset and information generated is not necessarily explicit; consequently the path from the goal to the solution is unknown (Simon, 1973) and as a result there is necessarily a large amount of problem solving.

A similar philosophy is presented by Rittel and Webber (1973). They describe problems as 'wicked' where; it is not clear when they have been solved, the information required to solve the problem depends on the proposed solution, the solution is good or bad rather than true or false. Wicked problems are symptoms of another problem and are also unique. Linking the arguments of Rittel and Webber (1973) back to Ullman (2002) it follows that engineering design problems are analogous to complex or '*wicked problems*'. Problems of engineering analysis or the solving of mathematical equations are seen to be '*non-wicked problems*'.

### **2.1.1. Problem Solving Processes**

Given that engineering design can be considered as a problem solving process (Pahl and Beitz, 1996; Ullman, 2002; Shigley, et al., 2003), a natural progression of design research has been to describe and prescribe the stages of design as a problem solving process. Pahl and Beitz (1996) strongly support the philosophy of engineering design as a problem solving process, and argue that design problems can be broken

down into smaller sub-problems and tackled systematically. Not dissimilarly from Simon (1969) they characterise problems as having three components. Firstly an initial problem state must exist, where the situation is undesirable. Secondly there should then be a desirable state or outcome to aspire to. The final prerequisite for problems is that there should be some barrier from achieving the transition from one state to the other.

Pahl and Beitz go on to suggest that there are several phases in the process of successful problem solving. Initially the goals and ideal end state should be identified, even if this is unrealistic (Pahl, et al., 2007). The next phase is to determine the boundaries of the problem and the constraints placed upon possible solutions. Removing prejudice as to the feasibility of solutions the designer or design team should create and develop a wide range of solutions. Having completed this, the solutions should be evaluated with respect to the criteria in the specification which should reflect the goals and constraints. The outcome of the evaluation stage can then be used to make decisions as to the chosen solution.

Overlaying the phases of problem solving, Pahl and Beitz detail some of the activities involved, Figure 2-1. Having confronted and identified the problem, information is required both with respect to the goals and constraints. Further gathering of information, definition, creation, evaluation and decisions all occur iteratively throughout the problem solving process. The arrival at a chosen solution occurs when a satisfactory outcome is reached.

An alternative approach to the problem solving methodology detailed by Pahl and Beitz (1996) is to use argumentation, breaking a problem down into small sub-problems and weighing up positive and negatives to each possible sub-solution.

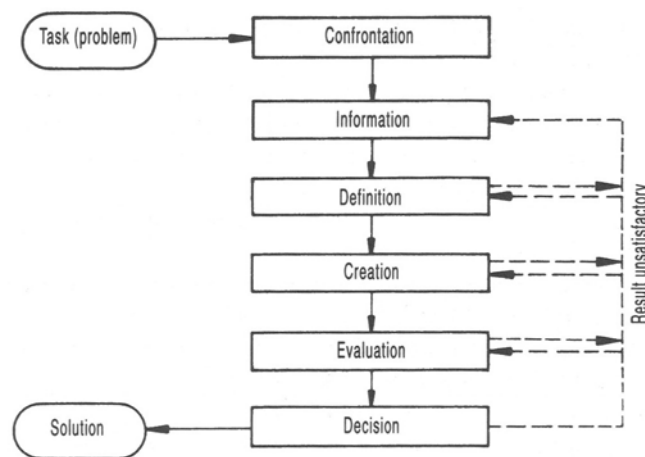


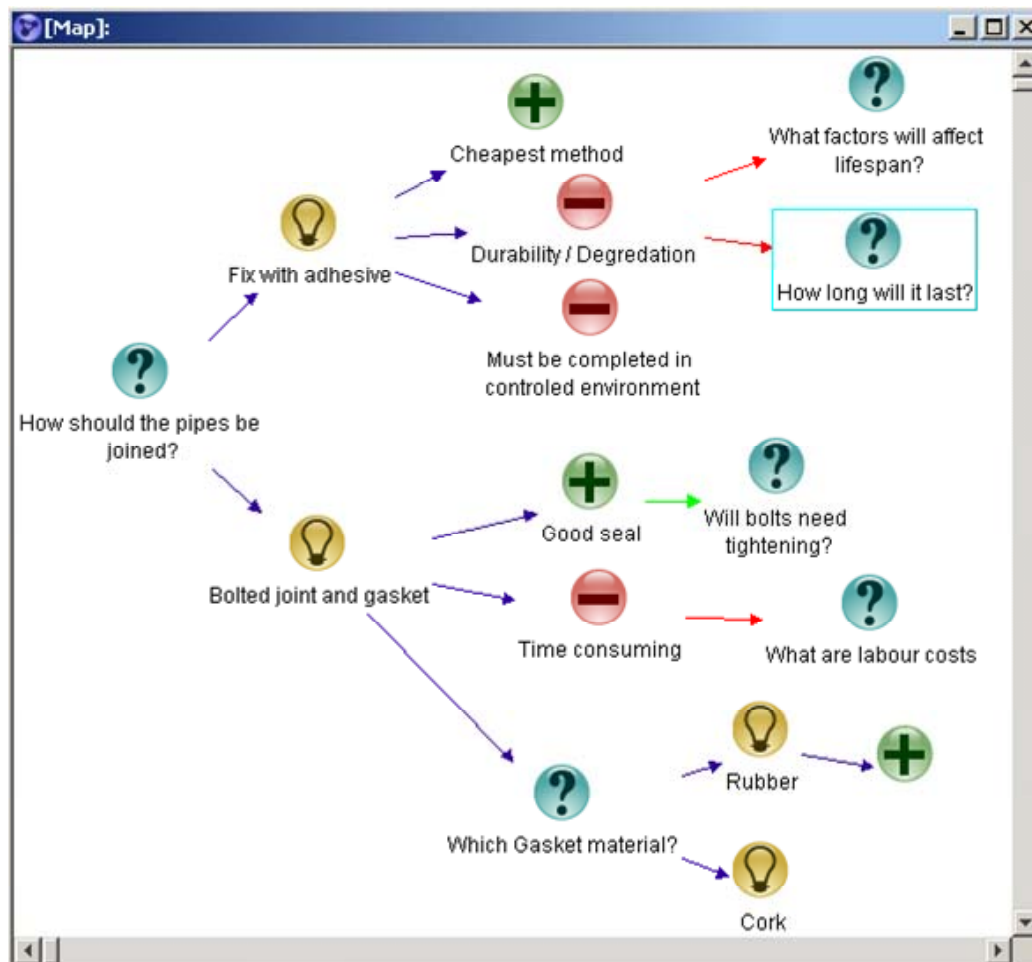
Figure 2-1: A solution generation process (Pahl and Beitz, 1996)

### 2.1.2. Argumentation Based Approach to Problem Solving

An alternative perspective on problem solving was created by Kunz and Rittel (1970) to solve complex or ‘wicked problems’. This argumentation based approach is named Issue Based Information Systems (IBIS). The original domain of application by Kunz and Rittel (1970) was in policy planning. Their work and argumentation based approaches have been extensively developed since then to cover general domains (Conklin, 2006) and engineering (Bracewell and Wallace, 2003). IBIS is most easily explained when it is applied to solve a problem by mapping argumentation. An explanation by way of example is now given.

As the name suggests the system is based on *issues*. These are identified and formed as questions, representing the problem or part of it. An issue would be “joining the two pipes” with the question being “how should the pipes be joined?” The second process is answering the questions by *proposing ideas* or solutions. In this case “fix the pipes with an adhesive.” *Arguments* for or against the solution should then be made; for example pros and cons might be, “adhesive is the cheapest fixing method,” but “adhesive will degrade more quickly than a bolt and gasket seal” respectively. When a truly ‘wicked problem’ is being solved, the arguments will identify further issues or questions which need to be answered. Based on weight of the arguments a *decision* can finally be

made. It follows that when IBIS is applied to record a problem solving process the rationale for decisions is captured by way of the argumentation. An example of applying IBIS to create an argumentation map is shown in Figure 2-2 using the software Compendium (Buckingham Shum, 2009).



**Figure 2-2: An IBIS style argumentation map produced in Compendium (Buckingham Shum, 2009)**

Designers and organisations will use a range of approaches to solve problems, be it argumentation maps or the Pahl and Beitz (1996) approach. Whatever method is used, the level and amount of problem solving will depend on the familiarity of the designer or design team to the particular design task, and is frequently expressed by the notion of maturity and originality.



### **2.1.3. Originality and Maturity of Design**

A number of the established engineering design texts draw distinctions between different types of design, more specifically the level of originality or maturity of the design. Based on the work of a number of authors (French, 1985; Pugh, 1991; Ulrich and Eppinger, 1995; Pahl and Beitz, 1996; Qian and Gero, 1996; Ullman, 2002) three distinct and commonly identified groups are presented, often referred to as Original, Adaptive and Variant design.

- An **Original Design** is one in which aspects of the design are not previously in existence (Ullman, 2002). This could be through the use of a new function or a new combination of exiting functions. Pahl and Beitz (1996) suggest these types of design are based purely upon the requirements and relationships between functions.
- **Adaptive** or **Redesigns** involve modification or development of existing products. The general structure, assemblies and components are better known than in original designs (Pahl and Beitz, 1996).
- **Variant** or **Routine** design elicits changes to the parameters of an existing design, the method for making these changes is well known, to the extent that it might be captured as a process or set of rules (Ullman, 2002).

As with any classification, completeness and distinction between categories is hard to achieve and the intricacies of the above three terms can be debated. Citing a design as Original, Adaptive or Variant therefore requires thoughtful argumentation. When observing design projects, it may be equally useful, and indeed simpler, to recognise how a design task reflects each of the categories, thus providing an understanding of the challenges faced by the design team. For example particular assemblies or sub assemblies might be original but the overall design

task might be adaptive. This is supported by the fact that the majority of design activity is recognised to comprise of either adaptive or variant design (Ullman, 2002). Although the prevalence of original design depends heavily on the knowledge and previous experience of the parties involved, for today's complex products and systems, the engineering project will necessarily involve multiple organisations.

#### **2.1.4. Projects and Organisations**

Designing rarely takes place as an isolated activity but is invariably the manifestation of a project supported by an organisation. The scale and complexity of engineering design projects can vary enormously, ranging from aircraft design, through to the development of a clutch for a car (Dandy, et al., 2008). Typical projects could fall into the domains of Aerospace, Automotive, Marine, Construction or Product design. Despite their great variety Smith (2007) identifies commonality in all engineering projects, having a degree of novelty and being an investment of resources with the objective of achieving some future outcome. As discussed in Chapter 1, projects often require people with a range of expertise to form interdisciplinary teams to achieve successful project outcomes. For these reasons amongst others, controlling and managing engineering projects is a complex activity in which effective communication, collaboration and information sharing are essential (Bouarfa and Mohamed, 2010). For this reason aspects of project management (PM), information management (IM) and communication are discussed further in sections 2.3, 2.5 and 2.6.

The notion that projects are supported by organisations highlights that they will be subject to influences from outside of the design environment. Organisations typically commission design projects where there is a commercial interest. In product design a classical distinction is made between two kinds of company; innovators and followers (Teece, 1986). Innovators invest highly in research and development to produce products which are new to the market and thus command a premium

price. Followers uptake these technologies and enter the market place later aiming to sell larger volumes albeit at a lower price. Their profit is achieved by redesigning to drive down the unit cost and taking advantage of the initial research expenditure afforded by the innovators. In this scenario the type of organisation will influence the choice of design projects, innovators tending for original designs whilst followers towards adaptive.

As discussed, various types of design activities involve a range of organisations and as a consequence the overall design and product lifecycle need to be managed to assure quality of design, in cost and on time.

## **2.2. Design Process and Product Lifecycle**

So far an overview of engineering design activity has been presented; but little has been mentioned of its temporal nature. The array of activities varies enormously through the evolution of a design project and it is important to understand these changes; this is especially the case because emails are used throughout the duration of a project. First the design process is discussed followed by some of the activities either side of this. Aspects of costing and product management through life are also discussed.

### **2.2.1. *The Design Process***

The nature of engineering design has already been described as an iterative problem solving process. Design research has explored the activities which comprise this and the procedures which designers follow, with the intent of improving the efficiency and effectiveness of the process. Widely accepted models include those by French (1985), Pugh (1991) Figure 2-5, Medland (1992), Ullman (2002) and Pahl & Beitz (1996) Figure 2-3. Although distinct from one another, all of the models reflect the four basic phases generally accepted to comprise the design

process: clarification and specification; conceptual design; selection and embodiment; and finally, detailed design. These models all share the inclusion of feedback loops showing the iteration in the process. These iterations are particularly important and necessary as understanding increases and generally result in modifications to the final artefact.

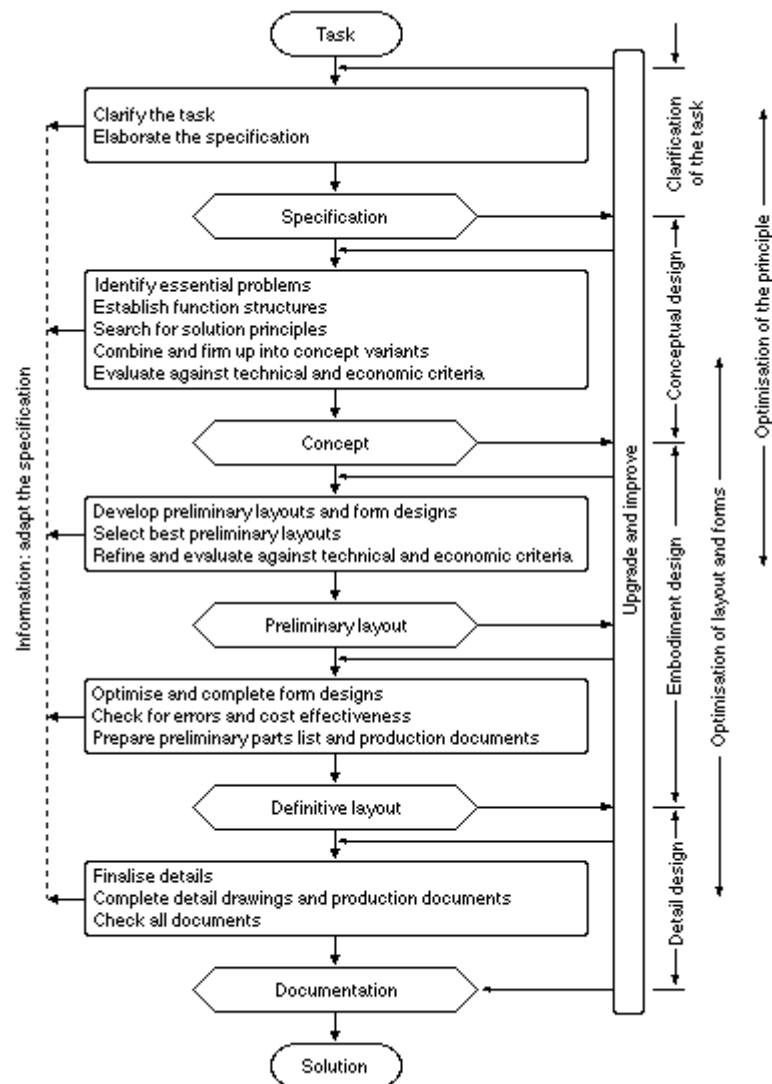


Figure 2-3: The design process (Pahl and Beitz, 1996 pg. 56)

### Step 1: Clarification and Specification

Although a small part of the process, a clear, thorough and precise understanding of the need for the design is essential to a successful project outcome (French, 1985). This stage involves determining the true requirements of the user or customer and hence setting the objectives that the final design wishes to achieve. Subsequently the constraints

upon possible designs need to be identified; these could include issues such as the availability of materials and financial resources, legislation, environmental considerations and technological limitations. With this achieved, a specification can be derived, providing criteria for which to evaluate design solutions throughout the process.

### **Step 2: Conceptual Design**

The next stage entails the generation and development of multiple design concepts to meet the design requirements, before selecting a final solution to proceed with. A key to success is seen to be the generation of a wide variety of ideas including those used in existing designs. Combining desirable characteristics from different solutions is commonly used to arrive at refined solutions and to further improve these. Some initial analysis as to the feasibility of design options is also conducted. Pugh (1991) emphasises the value in undertaking many iterations to generate the best possible solutions at this stage, as characteristics engineered into initial designs are more complex and expensive to change at a later stage. Having considered all of the available options with regards to the specification selection of a concept or concepts can then be made to take through to the embodiment design stage.

### **Step 3: Embodiment Design**

The embodiment stage converts design concepts into developed and tangible solutions. At this point the focus is on ensuring that the final design will perform as required. The layout and form of the finished artefact are set out and scaled and most of the standard components selected. Pahl and Beitz (1996) acknowledge that this stage requires many actions to be performed simultaneously; epitomising the iterative nature of engineering design and that there is no single correct starting point for analysis. At this stage the issues of design reliability and cost should be thoroughly investigated, along with manufacturing and assembly options.

#### **Step 4: Detail Design**

The detailed design phase concerns the optimisation and finalisation of every element of the design, to the point where complete manufacturing drawings and instructions are produced. Some authors (e.g. Pugh, 1991) treat detailed and embodiment design as one phase, due to the indistinct boundary between them. Even those who do make the distinction (e.g. Pahl and Beitz, 1996) describe detailed as the last stage in completing the design embodiment. The main difference between the phases is perhaps not the engineering activities, but the focus of them. Embodiment design seeks to ensure the design realises its basic requirements whereas detail design goes beyond this, seeking to optimise the solution and making it and its production as effective as possible. This is particularly challenging for today's complex multi domain products. As a consequence higher systems level approaches have been developed. Conceptually these approaches remain similar but deal also with the decomposition and integration of sub-systems.

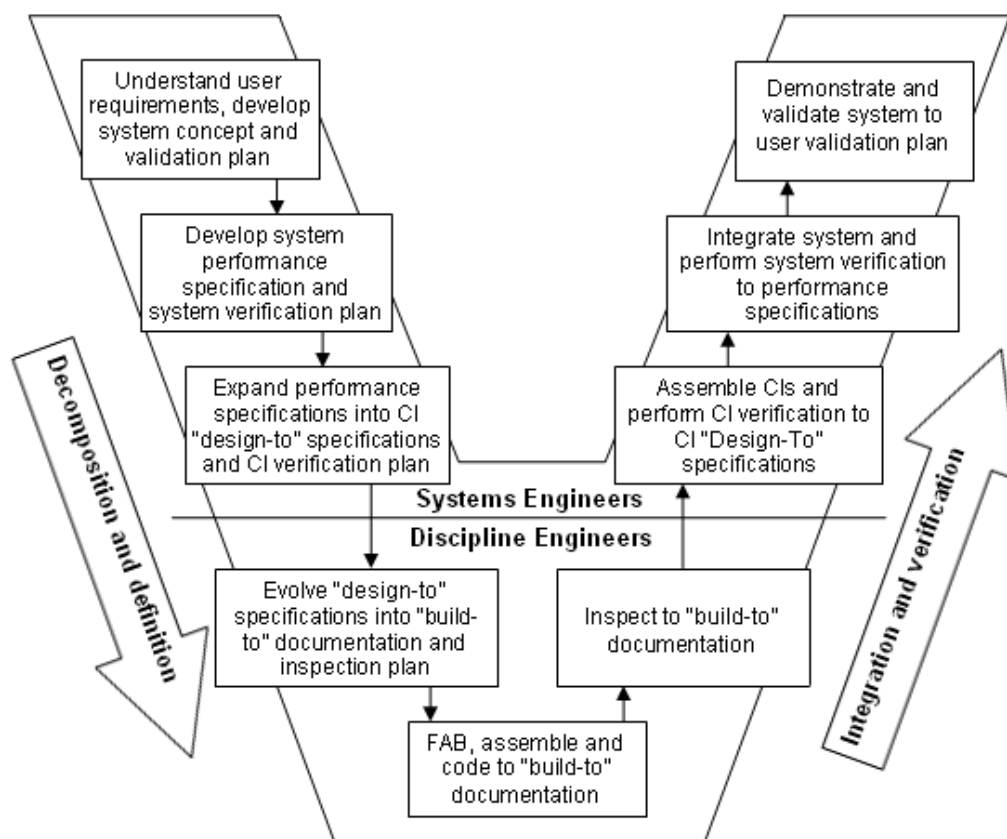
#### **2.2.2. Systems Engineering Design and Integration Process**

Systems engineering is defined as a ***“discipline that develops matches and trades off requirements, functions and alternative system resources to achieve a cost effective, life-cycle balanced product based upon the needs of the stakeholders”*** (Buede, 2009). Products conceived in this domain tend to be large and complex, such as a space shuttle, requiring the combination of many sub-systems which the traditional design processes do not include (Blanchard, 2004).

As an alternative the Vee model (Forsberg and Mooz, 1992) Figure 2-4, is widely referred to for the design and integration of such products (Aughenbaugh and Paredis, 2004; Blanchard, 2004; Buede, 2009). Elements of the model are similar to the traditional models of design but there is more emphasis on the specifications and testing required to produce an integrated final product (Blanchard, 2004).

The approach of the Vee model is to first decompose the overall system and define its requirements (left hand side of the Vee, of Figure 2-4). Sub-systems are then integrated to achieve this and the final performance verified (right hand side) (Forsberg and Mooz, 1992; Buede, 2009).

The first stage (working from the top left of Figure 2-4) is to gain an understanding of user requirements and develop a concept of the system. A plan is also made of how the final system can be validated against the user requirements. The following two stages then progressively develop a system specification to achieve this. This is similar to steps one and two (clarification and specification, and conceptual design) of traditional design process models.



**Figure 2-4: The Vee Process of System Design and Integration. Based on Forsberg and Mooz (1992).**

The initial three stages are undertaken by Systems engineers, the following three, at the bottom of the Vee are undertaken by engineers of

the appropriate discipline (Aughenbaugh and Paredis, 2004; Buede, 2009). In these stages the specification is further defined for the design of each sub-system and plans for the assembly and integration of these are made. These stages are similar to the embodiment and detailed design phases of the traditional design process models (Aughenbaugh and Paredis, 2004).

The final three stages are again undertaken by systems engineers. The first of these is the assembly of the discrete sub-systems, termed configuration items (CI's), these are each verified against their specification (Forsberg and Mooz, 1992). The following stage integrates these to form the system. The final stage is the validation of the system against the initial user requirements, based on the plan from the very first stage (Blanchard, 2004). This level of testing is not included as a core part of the traditional models of the design process.

For systems engineering the extensive testing and verification is a necessary extension to the traditional design process in order to handle the scale and complexity of systems and sub-systems. In the following section other wider aspects of the design and product development process are discussed, applying equally to systems engineering and product design.

### ***2.2.3. Beyond the Design Process***

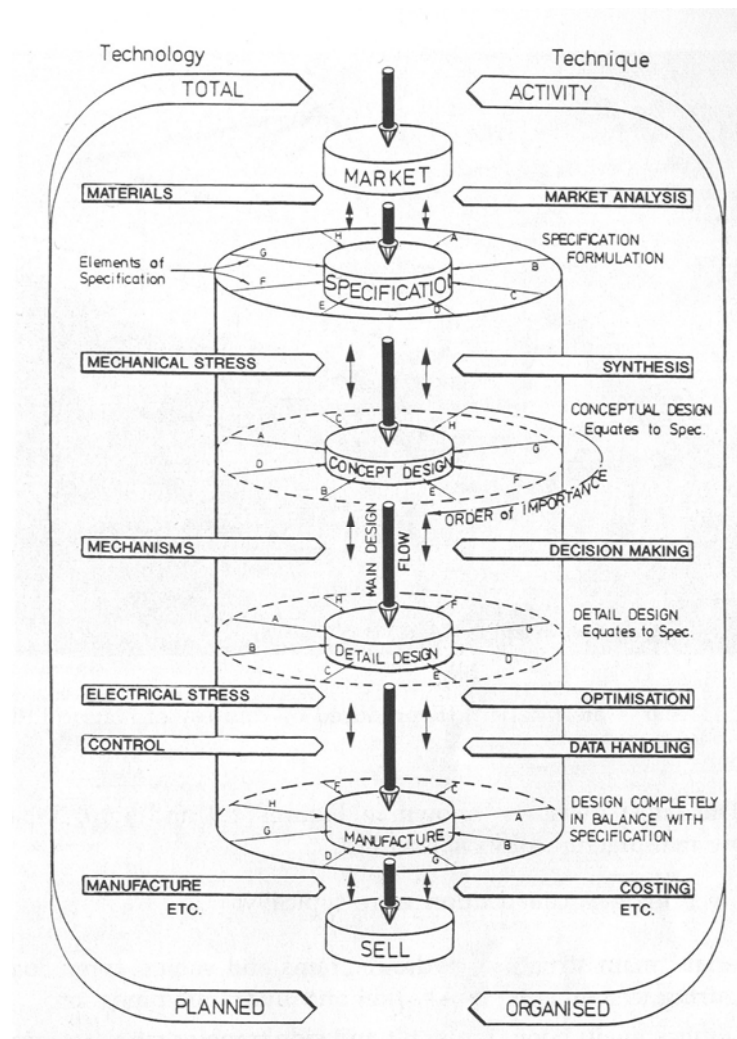
So far the core parts of the design process have been discussed, specification, conceptualisation, embodiment and detailed design, or in the case of the Vee model decomposition and definition through to integration and verification. From the view of the design project however, the preceding and proceeding activities are equally significant. One of the authors to highlight such activities is Pugh (1991), Figure 2-5.

Following detailed design, the product comes into being through manufacture. The consideration of this should already have taken place



throughout the design process, with the final planning part of the detailed design phase. Modifications may still need to take place during manufacturing, should a problem emerge. Furthermore information from this part of the process can be fed back into the design stages of future products, particularly newer versions.

Understanding the need for a design is described in the first part of the clarification and specification design phase. Pugh (1991) separates this further, determining that some form of market research is likely to take place which identifies the opportunity for a project and hence initiating the design process. The last stage of Pugh's process is the selling of the finished product, which he ties back with the market research stage that initiated the design process. Differences between mass, batch and one off productions are noted in relation to these.



**Figure 2-5: The total design process model (Pugh, 1991)**

In the case of a one off product the sale is likely to have taken place via tendering before the design process commenced. The specific relationship between the customer and designer should enable a correct specification to be developed producing the desired artefact. Should the design fail to achieve this, serious problems emerge. In the case of a batch produced product, initially commissioned by one customer, a tailored specification and pre-agreed sale should be possible as with one off products. There may also be an opportunity to sell further batches to other customers. For mass produced items, such as cars, the understanding of the market is at its most vital and complex. Usually each unit will be sold individually to an end user the design team will never meet. Thorough market research is a necessity and investment in marketing will be required to ensure sales are achieved.

The model presented by Pugh (1991) gives a holistic view of the design process, expanding on the core activities, and demonstrating the opportunities for information feedback throughout design. This forms an argument for organisations to observe and learn from their design projects. Since Pugh's work, the product models have been further expanded to encompass the extended product lifecycle which is now discussed.

#### ***2.2.4. Product In Service Support***

As previously stated the traditional product design models (e.g. French, 1985; Pahl and Beitz, 1996) have been extended to encompass the 'through life' phase (Pugh, 1991; Ullman, 2002). Early work by Ullman (2002) models the design process to incorporate product support, demonstrating some of the areas in which information generated in the design stages is required downstream during the product life and visa versa to improve design. This includes documentation, such as operation manuals and instructions for maintenance and decommissioning. It is also recognised that maintaining relationships with customers and suppliers, is part of longer term product support and might be useful for

initiating future projects. Furthermore knowledge as to the roles various parties played during a project may be useful particularly if in service problems emerge. When a project team disbands this kind of information is easily lost, hence all records need to be managed.

Where companies provide servicing for products in use they will incur costs at this late stage. These will be related to decisions made throughout the earlier stages of the design process. For this reason companies need to understand where costs are expended and incurred throughout the development of a product and also over its lifecycle.

### **2.2.5. Costing Over Product Lifecycles**

The design process models and discussion so far have already eluded that changes later in the design process bear a higher cost than changes in the earlier phases. Various studies have investigated the nature of this in relation to different design types and domains (Rush and Rajkumar, 2000; Ullman, 2002).

A general example for mechanical design is given by Ullman (2002), Figure 2-6, showing committed cost and incurred cost. It is recognised that a large proportion (75%) of the finished and operating costs of a product are determined by decisions made early in the design process (especially conceptual design), at this point the costs become committed (Newnes, et al., 2008). The actual expenditure, or point at which cost is *incurred*, only occurs after the detailed design phase when the manufacturing equipment is set up or (in the case of building) construction.

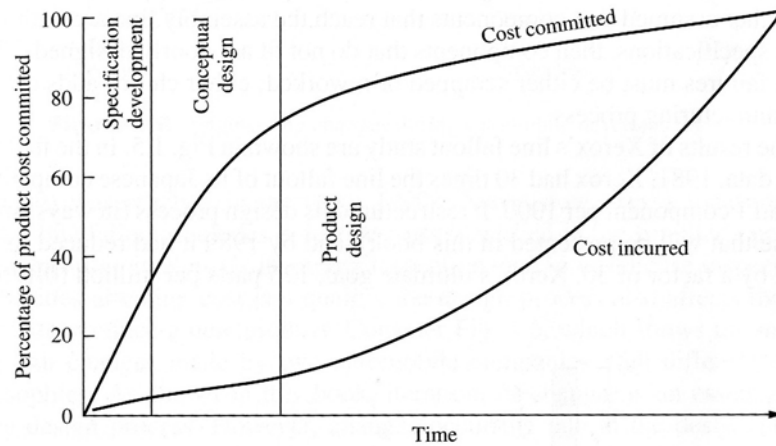


Figure 2-6: Commitment of manufacturing cost during design (Ullman, 2002).

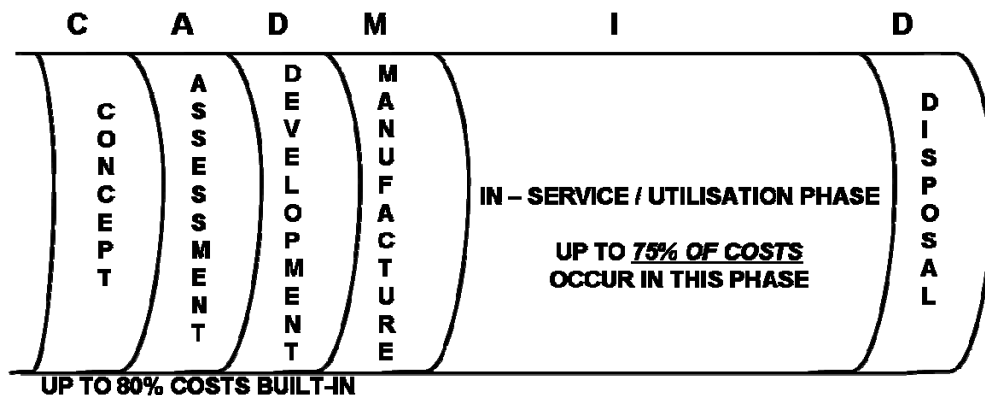


Figure 2-7: Commitment of costs and occurrence of costs in the CADMID Lifecycle (Ministry of Defence, 2009)

A similar pattern is seen for product service systems. This is illustrated by the CADMID lifecycle used by the Ministry of Defence (2009), which shows that up to 80% of costs are committed by the end of manufacture but 75% of costs are incurred in the in service or utilisation phase, Figure 2-7.

The issue of early cost commitment emphasises the importance of decision making in the initial stages of the design process. The ability to inform this by understanding how costs evolved during previous projects and how they behave in service is crucial to improvement (Cheung, et al., 2007). This presents a strong argument for organisations to learn from

previous design projects and the importance of records alongside effective communication and collaboration.

The stages of cost commitment and cost occurrence follow the lifecycle of a product, from conception through to disposal. Hence, effective management of a product's lifecycle is of paramount importance.

#### **2.2.6. *Product Lifecycle Management***

The term Product Lifecycle is used to describe various facets of a product life from a number of perspectives (Stark, 2004). Marketing lifecycles include the growth, maturity and decline in sales. Environmental resource lifecycles consider the extraction of raw materials and conversion into products, the impact of their use and finally their return to the environment by reuse, recycling, energy recovery or disposal. From the perspective of the manufacturer the lifecycle contains the design processes previously discussed including the wider activities such as product service and support (Saaksvuori and Immonen, 2008). From a user's perspective Stark (2004) describes five phases in this lifecycle: imagination (e.g. desiring a car), definition (knowing what car it will be), realisation (acquiring it), use (owning and driving it) and disposal (selling or scrapping it).

Understanding the complexities of product lifecycles ultimately enables companies to improve their products and services. This is increasingly achieved through the methodology of Product Lifecycle Management (PLM) which integrates activates, processes and systems throughout the phases of product lives. Central to the effectiveness of these systems is the capture and management of information and communication (Prasad, 1995). PLM systems are a key tool in the management of engineering projects, particularly those with long lives. Other aspects such as resources, scheduling, project risks and contracts must also be considered.

## **2.3. Managing Engineering Projects**

It was stated earlier (2.1.4) that all engineering design takes place within the domain of a project, and that it is usually undertaken with the objective of producing some profit or benefit to an organisation. It follows that effective engineering project management (PM) is a central element to achieving this. For this reason, the following sections present a fundamental background to PM. The ideas discussed are generally well supported in the literature, rather than state of the art research, the texts of Smith (2007), Gray and Larson (2000), Hales and Gooch (2004), and Field and Keller (1998) are predominantly drawn upon. The objective when organising projects is to make the most efficient use of time and resources and ensure obligations to all parties are satisfied (Smith, 2007). It follows that the core aspects to PM include scheduling, resource control and risks, and how these can be brought together and controlled through contracts. The discussion emphasises the importance of information sharing and communication.

### **2.3.1. Scheduling and Resources**

Engineering design projects require the undertaking of a great array of interrelated activities, the complexity of which only grows with the size of the project. A critical part of any project plan is scheduling of activities which is inextricably linked to available resources (Faulconbridge and Ryan, 2002).

A number of different resource types are differentiated between, including human, physical, financial and documentary (Hales and Gooch, 2004; Smith, 2007). The term human resources is used in relation to designers, technicians, administrators or any person able to fulfil a role within the project. Not only is their availability of concern, but also their expertise and abilities, along with the cost of their time. Various *physical resources* are required throughout a design project (Kerzner, 2006). During the design stages these might include simulation equipment, rapid prototyping machinery through to the basics of computer equipment and

relevant software. During manufacturing stages factory space, machines and tools will all be required. The *financial resources* available to a project can greatly affect its outcomes. The availability of capital and the willingness to invest it in new equipment or tooling might, for example, present further options to a design team. *Documentary or Knowledge Resources* capture a project team's ability to carry out new tasks based on previous experiences. Processes or techniques previously applied might be recorded in a documentary form. Team members might know and understand characteristics of certain design problems from experiences of previous design projects in similar domains (Schön, 1983; Oxman, 1994; 1996).

The availability of necessary resources and the dependencies between various activities affects the sequence in which they can occur (Kerzner, 2006). This is extenuated during manufacturing or construction phases, although applicable also during design. Changes and uncertainty in the process must be anticipated and schedules should maximise flexibility for this. With so many factors to consider, network algorithms are often used to help optimise work flow in conjunction with subjective understanding based on experience (Gray and Larson, 2000). Schedules can be directed to minimise timescales or maximise efficient use of resources; although a solution which optimises both might often be sought (Field and Keller, 1998). For design activities where problems are ill-understood such advance scheduling is not possible. As a result schedules are generated dynamically through processes of communication and collaboration.

### **2.3.2. Risk**

Ideally a project will deliver the correct solution, in time and on cost. Risks pose a potential barrier to achieving this, and can be considered as a combination of the likelihood of an undesirable event happening and the consequences if this does occur (Gray and Larson, 2000; Kendrick, 2003). As risks become realities, extra costs and longer schedules are

incurred and in the worst case projects can ultimately fail. Successfully managing risks then requires achieving a reduction in the levels of uncertainty associated, such that they are acceptable to the parties, and particularly investors, concerned (Wang and Roush, 2000).

The nature of different risks varies greatly, Smith (2007) distinguishes between those which are global risks and those which are elemental risks. Global risks are likely to affect all similar projects and might include political, legal, commercial and environmental influences; they are naturally harder to control than elemental risks but to some extent easier to predict. The elemental risks in a project regard the successful implementation and operation of the engineering design process and subsequent manufacture, as well as securing the necessary financial revenue from the outcomes.

Smith (2007) also classifies risks as commercial and technical risks, transcending the boundary of global and elemental risks. The commercial risk in a project essentially concerns its financial viability (Kendrick, 2003). Balancing expenditure against revenue is central to managing this; understanding influences upon these, thus enabling accurate predictions, will help to minimise risk. In some cases higher risk ventures may be deemed acceptable to an investor or company, as other perceived commercial benefits, such as competitive advantage, may be seen to exist. Projects are also fallible to technical risks, the result of which is the artefact not performing as it intended (Wang and Roush, 2000). It may sometimes be possible to correct this by the further investment of time and money however this is likely to increase the commercial risks. Technical project risk is recognised as being higher where the technologies are less well known to the designers, organisation or wider engineering community. Hence original design projects are usually associated with a higher level of technical risk (Smith, 2007).

An effective way of managing both commercial and technical risks in projects is through the use of reviews at various stages in the design



process. These can be used to highlight issues of risk so that they can be dealt with early and also to terminate projects deemed to incur excessive risk (Kahn, 2005). The early highlighting of risks away from the formal review process is therefore highly desirable. Such issues are likely to be implicitly represented in communication which is not formally recorded and analysed, such communication mechanisms include video conferencing, telephone and email.

### **2.3.3. Contracts**

Invariably a project involves commercial partners and outside organisations to a lesser or greater extent in the various stages of the design and manufacturing process. In some projects there may only be a customer and a raw materials supplier, others will involve hundreds of suppliers of critical sub-systems, materials and multiple end users. For example 65% to 70% of a Boeing aircraft is produced by outside sources (Destefani, 2004), although this figure may reduce it is still likely to be significant (Webber, 2009). Contracts are a two way mechanism of formalising relationships between these parties, recognised in law, and any written communication can be interpreted with such status.

In PM, contracts are seen as critical in controlling risks and setting out rules so that expectations are met (Carmichael, 2000). Obligations of the various parties are stipulated and agreed, thus protecting others from the responsibility and cost incurred by a party failing to meet its obligations (Field and Keller, 1998). The objective of this rigorous process is to ultimately achieve a project which meets its expectations in terms of cost, time and quality. Unsurprisingly contracts can become incredibly complex and are often used strategically in the management of projects. The importance of the contracts to any engineering project cannot be overstated; failure to handle them correctly can lead to the downfall of a project or even organisation. Of particular concern to organisations are records for communication which can impact upon existing contracts (Genborg, 2005; Blum, 2007) For example a record of a conversation or

meeting where additional functionality or specification changes are agreed. The potential issues surrounding communication are becoming of increasing significance as projects and design teams become larger and more distributed. Effective communication is a prerequisite for design changes to be made quickly where parties are distributed, and is essential if work is to be outsourced (Webber, 2009). The management of engineering teams and the information they communicate is complex, particularly where they are distributed and multi-disciplinary.

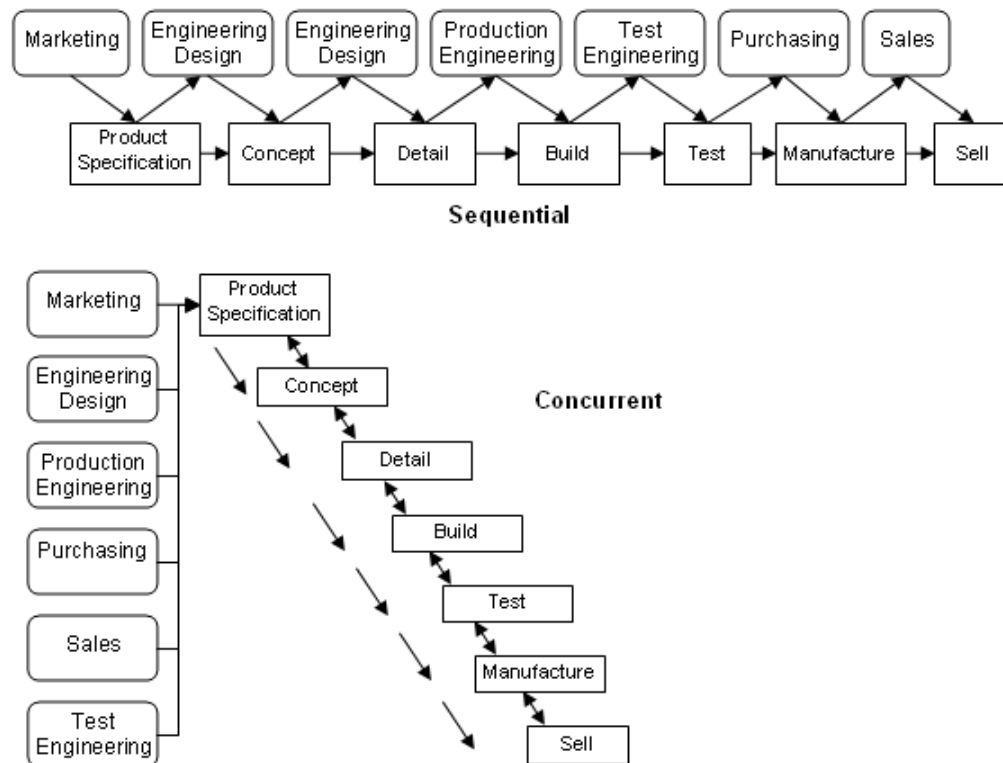
## **2.4. Design Project Teams**

Teams and team working are central to modern engineering design activity (Cross and Clayburn Cross, 1995; Gibson and Cohen, 2003). Not only are teams more effective at solving problems than individuals (Robson, 2002) but they are capable of achieving large scale projects requiring a broad expertise and thousands of hours of work. Project teams include intra-company members, such as administrators, technicians, engineers and managers and inter-company involving consultants, suppliers, contractors and customers (Brereton, et al., 1996). Projects teams can vary enormously but in particular, modern engineering teams are often large, interdisciplinary, globally distributed and virtually connected. Communicative and successful teams are key to successful projects (Jarvenpaa and Leidner, 1999).

### **2.4.1. Large Inter/Multi disciplinary Teams**

The concurrent engineering methodology overlaps phases in the design process which would traditionally have run sequentially. Early on in the concurrent engineering process consideration is extended to the later stages, particularly manufacturing Figure 2-8. Along with reducing lead times, this can help in achieving continuity and reducing the number of changes required in the later stages of design thus reducing costs. Whereas previously design information was passed “over the wall” from one team to another, now there is increased interaction between parties

from different phases of the process (Backhouse, et al., 1996; Tomiyama, 1998). Manufacturing engineers, for example, may work with designers in the conceptual stages. Team working across boundaries is therefore central to achieving the concurrent engineering philosophy and as a result communication is a fundamental component of concurrent engineering (Prasad, 1995).



**Figure 2-8: Concurrent and sequential engineering processes (Backhouse, et al., 1996)**

Many products are now so large and complex that no single individual could have the time or expertise to tackle them alone. An aircraft will contain hundreds of thousands of parts requiring tens of thousands of engineers to develop (Eckert, et al., 2005). Consequently team working is essential in engineering design projects. From large complex artefacts through to household items, many products integrate technologies from different domains; particularly electronics, software and mechanisms (electromechanical). The development of these products necessitates team working between the respective engineering disciplines during the

design process. Teams designing at these intersections are often referred to as interdisciplinary. Project teams are often formed by diverse groups of people and disciplines (Newnes, et al., 2006), their interactions (including information sharing and communicating) are key to project success and a core part of modern engineering design. However enabling such communication and collaboration is more complicated where teams are globally distributed.

#### **2.4.2. Global and Virtual Teams**

Increased global working between companies has promoted the adoption of virtual teams (Gibson and Cohen, 2003). These can be formed by individuals or smaller teams, either internally to the project or externally, to involve contractors or suppliers. The parties are physically separated, but linked through various electronic communication media, such as email or video conferencing (Kahn, 2005). Whilst the technical aspects of such collaboration are relatively simple to overcome it is recognised that Global Virtual Teams have different needs than those which are co-located, posing challenges to make them effective co-operators and collaborators (Massey, et al., 2003). The issue of building and maintaining trust within virtual teams is at the forefront of such challenges, and the benefits to teams of *face-to-face* introductions where possible are recognised (Jarvenpaa and Leidner, 1999). Other challenges include the transcending of cultural boundaries (when working globally) and that team dynamics can be more complicated to handle. These factors, along with varying organisational cultures, and differing objectives and goals of teams can further reduce effectiveness (Robson, 2002). Clear, timely and effective communication is central to addressing this. Of note is also Robson's observation that virtual teams are not as successful at tasks where high levels of creativity are required. Along with effective communication good leadership is important for the success of teams.

### 2.4.3. Leading and Organising Teams

Three types of team structure, Functional, Project and Matrix are commonly distinguished between (Bartlett and Ghosal, 1990; Ulrich and Eppinger, 1995), although it is recognised that few real teams will fall neatly into one category or another (Field and Keller, 1998). Engineering teams will tend to be Project based, Figure 2-9, (formed exclusively for the project) or Matrix based, Figure 2-10, where members have accountability to a project but also a division such as design or research and development.

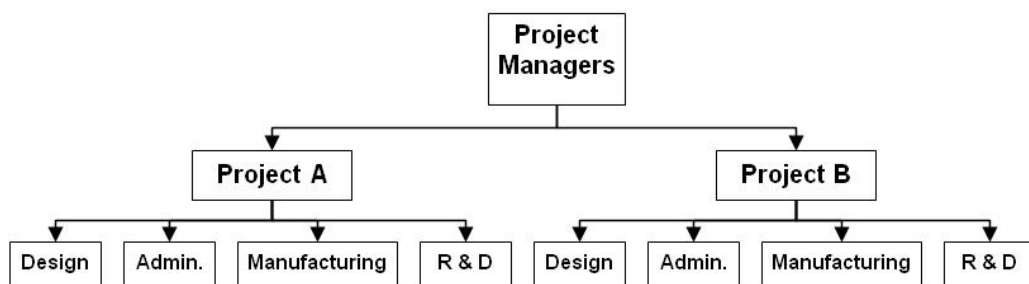


Figure 2-9: Project Style Team Structure.

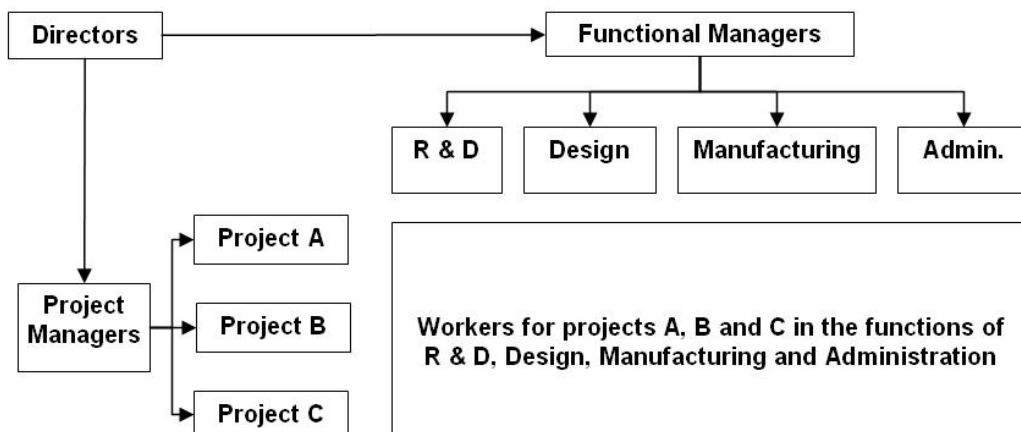


Figure 2-10: Matrix Style Team Structure

The structure is useful for establishing how lines of communication should flow and it is important that these structures enable and support this (Thamhain, 2003). The other aspect of a project structure is that it supports and details the relationships between team members in terms of their accountabilities and responsibilities. Heading the team will be the Project Manager whose effective leadership is historically seen as key in

enabling a successful team (Belbin, 1981). A balance of roles and necessary expertise is also recognised as important in forming effective teams and attention has been given to this in the context of engineering design teams (Wasiak, et al., 2008).

The ability of project teams to be adaptive is also useful in engineering design projects, where to some extent the solution and therefore path of the project is not known in advance. Increasingly self-directed teams with flatter hierarchies are emerging, which are more dynamic and versatile (Kahn, 2005). To balance project demands, thus improving efficiency, team size can be varied throughout the project. The loading of engineers in the earlier (particularly conceptual) phases of the project is often used to reduce lead times (Smith, 2007).

The keen participation of team members is vital in order to realise innovation (Thamhain, 2003). Senge (2006) contends that teams must have shared goals and objectives in order to be truly successful. Without such alignment members can be empowered but without having a positive effect. In conclusion successful teams are cohesive and focused. They should be well managed and selected and need to be flexible, but perhaps most importantly need to communicate effectively.

## **2.5. Communication in Design Teams**

The process of communication not only enables the transfer of information but also facilitates its generation via interpersonal interactions. Effective, timely, accurate communication of knowledge and information is a fundamental part of the design process. The theoretical background and definition of communication is a complex issue; which Eckert et al., (2005) explore thoroughly in the context of engineering design. It is not necessary to enter into such detailed discussions here, but based on their work, the following characteristics of communication are given.

- Communication concerns the transmission of facts or information.
- Communication processes are social ones, orientated towards other people.
- Communication is dynamic and temporal; contextualised by its time frame and relating to other events.

The way in which these different characteristics are enacted and their relative importance varies between communication mediums and the scenarios during which communication occurs.

### 2.5.1. *Communication mediums and synchronicity*

Communication can be achieved through any combination of visual, audible and tactile means (as noted by Huet, 2006). Mediums which enable these include *face-to-face*, emails, telephone conversations, facsimile, videoconferencing and instant messaging. The situations for these occurring also vary; for example *face-to-face* conversations could arise from informal chance encounters and ad hoc discussion, to highly structured meetings or interviews. A distinction is commonly drawn between these media is whether they enable synchronous or asynchronous communication, Figure 2-11. *Face-to-face* and telephone are examples of the former, where communication is instantaneous. Letters and faxes are examples of the latter, where there is a delay in which either party can act or reflect before replying. It can be argued that electronic text based communication mediums, email, instant messaging and forums can behave both synchronously and asynchronously (Shirani, et al., 1999), hence Figure 2-11 presents these mediums on a spectrum.

Asynchronous		←	→	Synchronous	
Fax				Face to face	
Letter				Web conferencing	
	Email			Telephone	
			Instant Messaging		
			Forums or message boards		

**Figure 2-11: Spectrum of asynchronous and synchronous communication methods**

In most studies emails are considered to be asynchronous because there is most often a delay of some minutes, hours or days between replies (e.g. Shirani, et al., 1999; Honeycutt, 2001; Genevieve, 2006). However they may be used by parties undertaking a rapid exchange of brief emails as if *face-to-face* and thus demonstrate synchronous traits. As the name implies instant messaging has synchronous traits, yet a delay in reply because a user is away from their screen renders the medium asynchronous.

There are benefits to communicating asynchronously via email. For example reviewing messages before sending cannot be achieved in face-to-face conversation (Ijsselsteijn, 2003). Synchronous communication is not always practical for both parties (Markopoulos, et al., 2003) such as communicating with someone in a different time zone, yet asynchronous email communication supports this. Conversely attempts to use email to communicate synchronously may be more time consuming than simply picking up the telephone (Day, 2009).

It follows that affordances of different communication mediums are based largely upon their variance in synchronicity, visual, audio and tactile channels. Video conferencing allows expressions, gestures and tone of voice to be carried; emails allow the recipient to re-read and make certain they understand before replying. Which communication medium is preferred will be largely situational, but it is useful to recognise the features of email compared with other mediums.

Research in the design community has highlighted that reference to objects either conceptual, such as drawings and sketches, or physical artefacts is an important part of communication (Eckert and Boujut, 2003). Interestingly it has been shown that email supports this functionality, and is more than a poor substitute for *face-to-face*, the use of attachments partly enables this (Ducheneaut and Bellotti, 2001).



### **2.5.2. Communication Scenarios**

A great deal of attention has been paid to the processes, purposes and acts of communication from both sociological and engineering research domains. These have examined from the fundamental components of speech (Bach and Harnish, 1979) up to complete interactions. Some of these ideas, particularly the more granular, are considered in more detail in Chapter 5 where the development of an approach to analyse email communications is presented. Three recognised high level scenarios for design communication are presented here; handover, joint design and interface negotiation (Eckert, et al., 2005).

Handovers occur when one person passes over their work to another, thus enabling the following activities to proceed and the design to progress. This 'over the wall' style of transaction is one directional, not producing any feedback. Despite the uptake of concurrent engineering methodologies this action is still recognised as being common (Eckert, et al., 2005).

During joint design scenarios, participants work synchronously and are usually co-located. The task tends to concern one problem, which the team work together to solve with discussion and instantaneous feedback.

In interface negotiations concern communication between participants whose tasks which mutually related to one another but are not sequential. Consulting with one another they must reach consensus on decisions affecting all of them; typically they will share estimated values of parameters. The objective is to make their elements of the design consistent with one another.

Some of the key processes at the core of these scenarios are seen to be idea generation, conflict resolution, decision making, requesting information, and making justifications. These shall be returned to again in Chapter 5.

So far this section has identified different communication mediums and communication scenarios which they support. This raises the question of what is currently known about the communication medium of email in an engineering design context.

### **2.5.3. Communication via Email**

There have been only a handful of previous studies regarding email in specific relation to engineering design. These have considered aspects of expertise location to improve knowledge sharing (Kim, 2002) and annotation of design rationale (Kato, et al., 2002). No studies however have investigated the general characteristics of email use in design projects. Although a particular niche, at the intersection of a number of areas, this is still somewhat surprising, given the prevalence of email use in engineering organisations. More attention however has been paid to email use in general contexts and this work and its applicability to the engineering domain shall be discussed in the following chapter.

A product of email communication is the generation of vast amounts of information, some of which would not have been captured if communication had taken place via another medium. This content, alongside other artefacts, such as reports and drawings, forms part of a project's information record and the design record. Hence managing this information effectively is important in both the short-term and long-term.

## **2.6. Information Management**

The creation and use of information has long been recognised as central to the process of designing and all collaborative activities. Newell and Simon (1972, pg. 788) propose that ***“when engaged in problem solving...humans are representable as information processing systems.”*** a sentiment reflected by Ullman (1992) who describes design as ***“the technical and social evolution of information, punctuated by decision making.”***

This section discusses the context in which information is important to engineering design projects and gives an overview of the processes involved in managing such information. First however it is useful to clarify what is meant by 'information' and its relation to the terms 'data' and 'knowledge' as discussed in Chapter 1. Informed by the works of Hicks et al., (2002) McAlpine (2007) and Lowe (2002) (who each discuss the terms origins extensively) the following definitions are given for the purpose of this work.

**Data** consists of characters, symbols and numbers, without context and in a way which is meaningless to a human. It will however typically be structured in such a way that it is machine readable.

**Information** can be formed from data, descriptions and context, and can be interpreted by a person. Diagrams and text are both examples of information. A further distinction is made between informal and formal information although this might be better viewed as a spectrum. Formal sources of information might include CAD drawings, reports or published documents. Characteristically these contain some degree of structure and adopt common templates or standards. They are often created for the purpose of communicating certain information impersonally but to a known demographic. Informal sources of information might include personal notes, emails, sketches and conversations. These typically contain little structure and can be regarded as personal information; in some cases those other than the creator or intended recipient may struggle to fully understand their meaning without additional context.

**Knowledge** is a personal understanding or belief. Individuals obtain new knowledge by the process of integrating information with their pre-existing knowledge. To transfer knowledge from one person to another it must be represented as information.

Having made the distinction between these three assets, it follows to discuss the management of these assets. An explanation of the concepts

of data management, information management and knowledge management are given for the purpose of this work. It is noted that these are interdependent and (as for the terms above) hard to delineate precisely.

The term **data management** is not widely used, but might easily be considered as a sub set of IM concerned with the digital level. Prevalent issues include, managing access rights to shared data, compatibility of software, preservation of archives and security. Such issues are discussed extensively within the digital curation community, and specifically regarding emails (Pennock, 2006).

**Information Management** (IM) encompasses such activities as organising, accessing and sharing information. Standards such as the format and layout of documentation prescribed by an organisation are also encompassed. The objective of IM is to make most effective use of information resources.

The concept of **Knowledge Management** (KM) is the most ambiguous of the three, being that knowledge is not a quantifiable or physical asset. KM is considered here to involve managing the environment in which knowledge is created generated and shared. It is also recognised that because information is a prerequisite for knowledge processes IM will affect KM.

The three concepts are undoubtedly interrelated and will in practice impact on one another. The main concern of this work is IM with regards to emails; e.g. how they are laid out, shared, stored and retrieved. There is also however an element of recognising what affect this has on KM; e.g. sharing of expertise. Given therefore the importance of IM to this investigation, the following sections discuss the relationship of information in engineering design.

### **2.6.1. Information in Design**

The importance of information to the design process and more generally design projects cannot be overstated. It is widely acknowledged that engineering design can be regarded as a process of transferring information from one state to another (Hubka, 1988; Ullman, 1992; Hicks, et al., 2002). Information is required as an input to the design process and is subsequently generated throughout the design process (Court, 1998). Engineers use information for a wide variety of purposes, to help understand problems, generate solutions and present decisions (Kwasitsu, 2003). Lowe (2002) found that engineers required information throughout all the stages of the design process. More significantly the sources for this information were often personal collections, rather than company wide information, highlighting the weakness in IM approaches. Effective IM is therefore a prerequisite for maximising design performance.

### **2.6.2. Information Management Lifecycle**

Having explained the importance of information and IM to engineering design projects, the key processes involved are now discussed. This is achieved through the use of an information lifecycle; numerous such models have been presented focusing on various aspects from digital curation and storage capability (Tallon and Scannell, 2007; Digital Curation Centre, 2009; IBM, 2009) through to linking information processes (SIMS, 2003). The lifecycle considered for this discussion, presented in Figure 2-12, outlines the phases of: creation and capture, organisation and storage, search and retrieval and re-use or disposal. The following discussion outlines some of the computational and user aspects relevant to these key phases and with particular reference to engineering design.

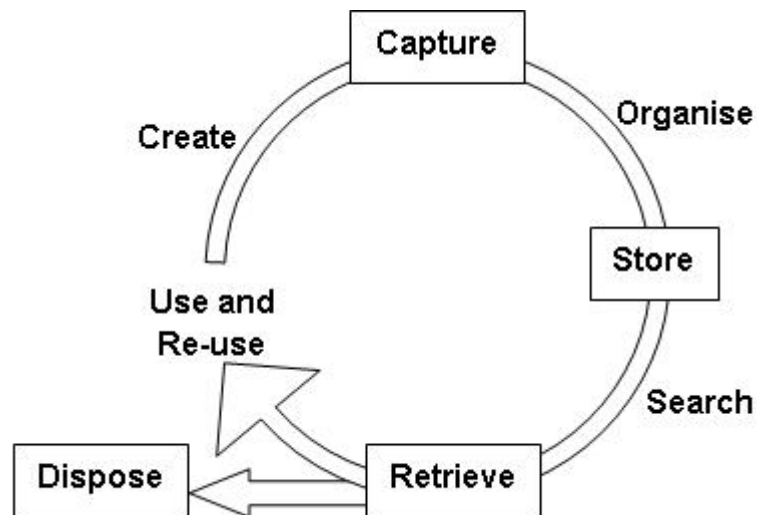


Figure 2-12: An information lifecycle, based on SIMS (2003)

### Creation and Capture

The first part of the information lifecycle presented is the generation of information and subsequent capture. Information is continuously generated throughout the course of the design project, represented in a wide variety of forms such as reports, CAD models, emails, notes, sketches, conversations and presentations. Information is always captured within such a context and not in isolation; furthermore part of this process is usually conversion to machine readable format (Hicks, et al., 2002).

The ease with which information is captured varies greatly between sources, as does the concision of the information. CAD models are created and exist within the computational environment; capture of this information is therefore essentially automatic (Regli and Cicirello, 2000). This is also true of articles such as reports and emails. At the opposite end of the spectrum are sources such as spontaneous informal work related conversations, although it is possible to record and capture these, there is a high overhead in doing so (Hindus and Schmandt, 1992; Whittaker, et al., 2008). Other information sources will fit in between these extremes, such as paper based documents which might be

digitised (Brown, et al., 2000), or formal meeting recordings and transcripts (Truong and Hayes, 2009).

A further difference between information sources is their content. It is the nature of a report to present a concise version of design discussion and explanation but as a consequence some points and a level of detail will be lost (Conway, et al., 2007). A paradox of this would be to record everything, for example a tape recording of a meeting. The level of detail is far more comprehensive if not complete, however useless information will be captured along with useful. This notion of 'information loss' is identified by Huet (2006) who explored the differences between meeting transcripts and meeting minutes and also McAlpine et al., (2006) with regards to engineers log books.

To make use of sources; such as recordings, personal notes and emails, at the later stages of the lifecycle a more intensive approach to filter and search may be required (McMahon, et al., 2004). Methods enabling the effective reuse of information from documents (Wild, et al., 2006), log books (McAlpine, 2010) and meetings (Huet, 2006) has been investigated. Email however has been largely overlooked. It is unique in that it is easily captured yet contains highly original and unabridged information and therefore provides extensive potential for reuse.

In addition to directly capturing information generated during design projects, there is a desire to capture further, often unarticulated, context to surround this. Specifically there has been considerable attention given to the process of capturing design rationale, the reasoning for decisions in the design process (e.g. Bracewell, et al., 2007; Giess, et al., 2007). One approach to achieving this is through the addition of supplementary marking up or tagging, discussed in the next sections. The alternative approach adopted by Bracewell (2007) is to generate and articulate rationale more explicitly as part of designing (consequently allowing capture).

## **Organising and Storing**

The organisation and management of information is critical for business success; especially design information in the engineering domain (McMahon, et al., 2005). Design projects can generate many thousands of electronic files of various types. When storing this information it is vital that files are organised or processed in a coherent way, to avoid chaos making any prospect of retrieval impossible. The term 'information overload' is used to describe this (Eppler and Mengis, 2004). Engineering organisations have typically adopted PDM and PLM systems to help achieve this. Both systems centrally store and organise files within a structure conducive to the project. PDM systems are focused particularly on compiling product modelling data, such as geometry, manufacturing and analysis data. PLM systems are more holistic, and include wider project documentation envisaged to support the product through life (Stark, 2004).

Despite the existence of such central management systems engineers also keep information on their personal computer space (Lowe, 2002). Furthermore Lowe found that this type of personal information accounts for around half or an engineers information usage. The way in which engineers manage this 'Personal Information' has also been closely observed (Boardman, 2004; Hicks, et al., 2008). Hicks et al., (2008) found that engineers tended to organise files and folders in hierarchies, which they argue represents a form of classification. Although the behaviour of individuals varied, there was a tendency to organise files in relation to product functions. File naming strategies were noted to be weak and particularly varied.

A further aspect in the organisation of information is the use of post processing to add contextual information to facilitate retrieval. Particularly in the case of informal information sources this can add considerable value, but can represent a significant investment. More established methods of adding contextual structure, particularly to documents, include



the addition of contents lists and indexes, sorting alphabetically or by date, and structuring with hierarchies (Dominich, 2008). It may be possible to post process emails but this is likely to be time consuming given the high volumes involved. This may however provide benefits in making them more searchable and their content more re-useable.

### **Search and Retrieval**

A number of authors have investigated information use by engineering designers and in particular the way in which they search for information. Designers have been found to spend between 20 - 30% of their time searching for information (Court, 1998). When doing so they have been shown to rely predominantly on human sources but also drawings, databases and reports (Auricchio, 2005). Engineers use people to find documents and documents to find people (Hertzum and Pejtersen, 2000). This somewhat blurs any division between IM and KM but acknowledges the importance of the individuals within either process.

Around half of information searches are targeted to retrieve factual information, one third to aid reasoning in design development, with the majority of other searches relating to deliberations, via analysis and evaluations (Auricchio, 2005). Sources included in such searches contain formal and informal information items within central systems and personal collections, both computationally recorded and by hand. Experts rely particularly heavily upon personal collections, and are prepared to search widely, mainly for text based information. In comparison novice engineers are more likely to search for geometric and CAD based data although will not go to such lengths to do so (Lowe, 2002).

Despite the importance of humans in locating information, searching electronic file collections is also vital to engineers. The domain of IR focuses on identifying and retrieving information from the computational environment to match the predetermined needs of a user (Meadow, et al., 2007). This has proved to be a non trivial problem which is far from

solved (Dominich, 2008). One facet of IR is the addition, formatting or structuring of information at the organisation and storage stage as discussed previously. The second facet is the application techniques, including algorithms and linguistics, to interrogate information from the computer environment. It may be that this second facet is applied to achieve the first autonomously. The computational techniques involved have been explored extensively (Dominich, 2008) and remain an area of high interest.

The fundamental difficulty with searching and querying techniques, particularly key word searching, is the limitation of artificial intelligence. Computers are not capable of meaningful cognitive understanding and can therefore only return results within the limitations of the algorithms written to achieve them. For example a computer searching for the articles containing the word “motorbike” will not retrieve articles where the word “motorcycle” is used instead.

A number of authors have proposed the use of taxonomies to aid retrieval, from both informal sources and from engineering documents, the generation of these is however highly labour intensive. Although computer generated taxonomies have been attempted (Yang, et al., 2005) they have not been found to produce suitably accurate results.

### **Reusing or Disposing**

The final phase in the information lifecycle described here is the reuse of information. Markus (2001) identified a number of different motivations for knowledge workers to reuse information. These included learning from experiences of previous and similar projects, novices seeking to gain expertise or further their knowledge, and to recall reasoning and decisions made during earlier stages of the current project. Reuse of information in engineering projects can influence the way in which design processes are undertaken and help develop best practices (Baxter , et al., 2007). The reuse of more traditional information relating to product

geometry and detailing, often supported by PDM systems, is also recognised as important (Hicks, et al., 2002). The re-use of design information has been shown to be highly useful even at the conceptual design stage, particularly when undertaking adaptive and variant design activities (Khadilkar and Stauffer, 1996). It is also suggested that re-using design information helps to minimise work and reduce project risks downstream.

Despite the benefits of reusing design information, designers have been found to exhibit a reluctance to do so (Busby, 1999). This social problem has been attributed to uncertainty of designers in when previous or existing work might be applicable to their new situation. The reuse of information will generally generate new information; capturing this, closes the lifecycle.

### ***2.6.3. Challenges for Information Management***

In the previous section was shown that information and IM are central to the undertaking of engineering design projects. As increasing amounts of information are generated, the effective organisation and storage of this becomes more challenging. Improved methods for sorting and distilling information are needed. This is particularly the case with a technology such as email, where large amounts of unstructured information covering a broad range of topics, purposes and varying significance are generated.

Encouraging the reuse of information by engineers may in itself be a challenge, which may partly be overcome by educating engineers of the benefits. More significant however is the problem of easily retrieving pertinent information, which undoubtedly discourages information reuse. More effective search techniques need to be developed, which can identify context, so that users can find useful information quickly, even if they have a limited understanding of what information exists or their own information needs.

## **2.7. Summary**

This chapter began by presenting a background to engineering design projects. Design and systems integration process models were identified (Forsberg and Mooz, 1992; Medland, 1992; Ullman, 2002), and aspects of project control including cost, risk and life-cycle management were discussed. It was identified that effective communication and information sharing is essential to support these processes (Hicks, et al., 2002; Bouarfa and Mohamed, 2010).

The trend towards distributed, large complex multidisciplinary teams (Jarvenpaa and Leidner, 1999; Gibson and Cohen, 2003) was identified. These large teams have arisen due to engineers designing increasingly complex products (Eckert, et al., 2005) and therefore often work with many partners with diverse expertise (Brereton, et al., 1996; Destefani, 2004; Newnes, et al., 2006).

This move results in a number of challenges to improve communication (Massey, et al., 2003; Webber, 2009), shared understanding and reduce knowledge loss (Huet, 2006). As such the role of information and communication in design projects and their importance to successful problem solving and project outcomes was emphasised (Simon, 1969; Hubka and Eder, 1982; Hicks, et al., 2002; Ullman, 2002).

A number of communication and information sources were contrasted such as phone, fax, engineering drawings, emails and minutes of meetings (Huet, 2006; Conway, et al., 2007). It was noted that information and communication activities vary throughout the course of project and product lifecycles (Eckert, et al., 2005).

It was identified that email is one of the few tools to be used throughout this project duration and is expected to be used to support a range of engineering activities. However, little research was evident in the area of email use in terms of design knowledge and project management.

Literature has illustrated that the use of email is common place (Bellotti, et al., 2003; AIIM International, 2006).

Overall it has been shown that communication is vital to the design process and that, as a widely used communication channel, email use is therefore also of significance. This view is supported by various industrial partners of the University of Bath<sup>4</sup>. There are two areas in which it is felt that the effectiveness of emails might be improved. Firstly, there is a need for more coherent management strategies for archiving, retrieval and reuse of information. These concern both direct use of email content and understanding of how organisational knowledge is referenced in email communication. Secondly, that general working process could be better supported particularly where users deal with large volumes of email. This also includes opportunities relating to improved collaboration and information management through to improved design and project processes. The contention is then reiterated that to make steps towards achieving these, it is necessary and desirable to form a substantiated understanding of email use in engineering design projects.

This chapter has partly addressed **Objective 1**; to examine the existing research and determine what is known about the role that email plays both in engineering design and other working environments. Literature was reviewed exploring email use in engineering design, (although little research has been conducted in this area) and exploring communication in engineering design. Chapter 3 continues by investigating the use of email as a communication tool in other working environments. It also provides a state-of-the-art review of email use within the working environment.

---

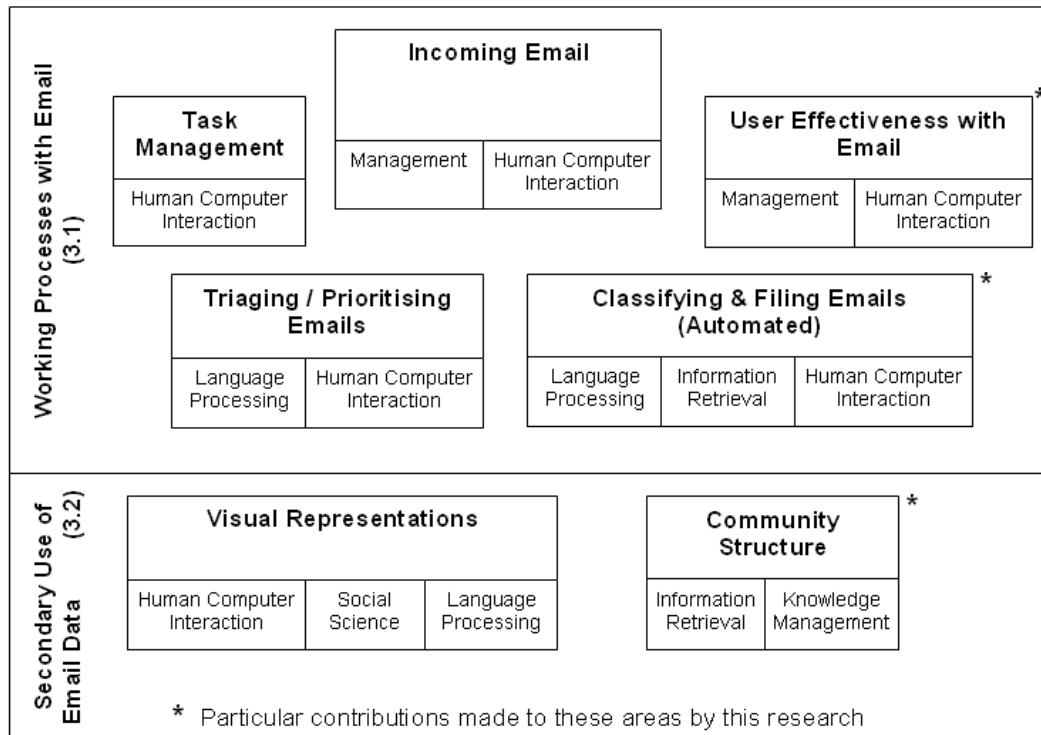
<sup>4</sup> Conversations with Project and Process Managers at Converteam, BMT Group and Airbus.

### 3.0 Email: State-of-the-Art

The findings from Chapter 2 highlighted the need to ascertain the use of email in engineering design projects, through considering other working environments. To investigate such use of email two specific objectives were identified:

**Objective 1:** *To examine existing research and determine what is known about the role that email plays both in engineering design and other working environments.* **Objective 2:** *To examine the tools and techniques proposed to analyse email use, including archiving and retrieval as well as general work processes.*

This chapter explores the wide variety of existing research which impacts upon email. This involves work that relates to users and their interactions with email and connected work processes, as well as the integration of Email as part of IM and KM strategies. (Broadly speaking the nature of research conducted encompasses both that which seeks to explore and analyse user behaviour, and that which upon the basis of this understanding, proposes tools to affect improved email use). Within this scope there are a wide range of applications and scenarios which research has addressed. Consequently the involvement of many communities and disciplines is represented in the following sections, these include; Human Computer Interaction, Information Retrieval, Language Processing, Management and Social Sciences. It is unsurprising therefore that division of topics within the existing research field for discussion is nontrivial, with overlapping ideas and many authors bridging fields. This space is shown conceptually in Figure 3-1.



**Figure 3-1: The relevance of various disciplines to topics of email research**

The first, and major section of this review (3.1), addresses common activities and ways in which email is used; both tools and user behaviour are discussed (grouped in the upper half of Figure 3-1). The second section (3.2) explores how email data can be post processed to reveal insights into both the records retained from communications and the communities which have created them (grouped in the lower half of Figure 3-1). The final section concludes, summarising key points from the review and their implications for this research.

The objectives of this research concern the above mentioned aspects of the working processes of dealing with email (3.1) and the secondary use of email data (3.2) as presented at the conclusion of this chapter in Figure 3-4. In particular however the outputs of this research inform the areas of: filing and classification, user effectiveness with email (in terms of appropriateness to task) and an understanding of the community structure within engineering design projects (highlighted in Figure 3-1.)

It is worth noting at this point that the review does not include discussion of junk, malicious and unsolicited email colloquially known as SPAM. Although a wealth of literature exists in this area it bears no specific relation to engineering design project email or the remit of this research.

### **3.1. The working process of dealing with email**

This section interlaces research addressing user behaviour, the nature of email communication and work flow with a selection of proposed tools and applications developed to improve email use. To begin the following sections one of the most influential and early studies into email use is presented. Building upon some of the ideas arising from this, the review then discusses tools and behaviour relating to four key email activities; receiving email, classifying and filing email, task management using email and triaging or prioritising email. This distinction of activities is similar to ones used in previous reviews, particularly Venolia et al., (2001). Finally some findings on the effectiveness of email for different processes and when compared to other communication media are presented.

#### **3.1.1. Whittaker and Sidner's Empirical Study**

Whittaker and Sidner's (1996) exploration of personal IM in email has established itself as one of, if not the, seminal work in this domain. Many of their original findings and principals are still relevant over a decade later. In their empirical study Whittaker and Sidner identified that the principle design of email clients intended messages to exist in two states: a small number of unread items in the inbox, and the rest, once responded to or simply read are filed appropriately. In practice this was not found to be the case, with large numbers of read messages in inboxes, and various degrees and styles of filing taking place. An explanation for this scenario gave rise to the term "email overload," referring to both the overloaded functionality of email as a tool and the volumes of email users receive.



Emails in many cases relate to a task the recipient needs to execute away from the email domain, as a consequence these messages are kept in the inbox acting as reminders of things to do; in effect the email environment became a task management tool. Users were also seen to keep messages relating to ongoing correspondence within their inboxes, acting as a reminder of what had been said, and awaiting reply following another process. It was also recognised that by virtue of creating a written record of communication, content of emails could be held to be legally binding. The fear of not being able to locate items once filed, and uncertainty about their importance at a future point in time led to users being reluctant to file emails and thus added to the inbox accumulation.

Along with the functions and processes already mentioned, document delivery, scheduling meetings, handling support queries and sending reminders were all found to be common place. Long messages awaiting reading and those whose status was unclear were also found to clutter inboxes.

In summary Whittaker and Sidner found that the inbox was the home of incomplete items for one reason or another. But it was the variety of functions for which email was used that caused such large volumes of messages to arrive, which users struggle to stay on top of. At the time of their study (1996) it is understandable that the perspective from which the domain was explored was one of filing, in line with the email clients of the time and more traditional office practices. It is perhaps therefore surprising that they recognised opportunities to improve email away from this traditional structure, with much of the following work (e.g. Venolia, et al., 2001; Gwizdka, 2002; Bellotti, et al., 2003) building on their recommendations and ideas. Although more tools are now available (Boardman, 2004) they are not necessarily utilised and the requirements identified by Whittaker and Sidner still exist (Fisher, et al., 2006).

### **3.1.2. Incoming Email...and too much of it**

Incoming email messages have been shown to affect work flow, with some users interrupting their activity to check each message as it arrives whilst others will dedicate time slots to email checking. Jackson and Burgess (2003) show that the reduction in productivity due to interrupt may have significant financial cost, and proposes that users check emails on an hourly basis to reduce this. The large volume of “send to all” emails were felt to contribute to this problem and it is suggested that these to be restricted within organisations (Jackson, et al., 2003).

Incoming email and its effects were considered by Venolia et al., (2001) as part of Microsoft’s investigation into email client requirements. It was identified that the significance of an incoming email to the recipient could be determined by factors such as the identity of the sender, whether the messages were direct or cc’d, the number of other persons the message was sent to and whether it was flagged as important. Furthermore it was recognised that it may be useful to have a more tailored notification system for message arrival, based upon these factors, with the possibility to display information in a side bar such as subject heading.

A more recent study of users email behaviour by Renaud et al., (2006) showed that more than 60% of individuals checked their email on a more than hourly basis. However it was reported that users found email to be less of an interruption than *face-to-face* communication and telephone calls; this is perhaps due to the integrated nature of email applications within the user’s working environment and an element of control over when and how to respond afforded by asynchrony. It was also found that the perception of the need for a quick response was felt more often by recipients than senders.

### **3.1.3. Automated Classification and Filing of Emails**

A large amount of attention has been afforded to the automatic classification of emails through computational techniques. Many of the

techniques are well established in the automated text analysis domain where detailed analysis and comparisons of them exist. With detailed discussion of these tools carried out extensively elsewhere (e.g. Strzalkowski, 1999; Cristianini, 2000) the purpose of this section is therefore to illustrate the variety of techniques available and identify how far research in this domain has progressed with respect to processing email.

Early work comparing *rule* versus Information Retrieval (*IR*) based classification techniques for email text was undertaken by Payne and Edwards (1997). IR methods examine biometrics of exemplar text already in folders and compare them to the email awaiting classification looking for similarities; rule based methods instead classify based on predetermined parameters such as the file size, sender or the frequency of key words. Payne and Edwards found that dependant on the sample data used either technique could perform superiorly. They also considered whether the use of metadata along with the email body aided filing. It was found that the inclusion of the subject line improved reliability with the rule based technique but had the inverse affect on the IR method. Both rule based and IR based systems have the ability to 'machine learn' (Marsland, 2009). It is worth noting that many of the tools were developed using emails which had already been classified, using a proportion as the training set and the rest as the test set. It is recognised that this does not account for incorrect filing in the original set, furthermore email is dynamic and tools need to be tested in a live environment (McCreath and Kay, 2003).

A strong example of a retrieval based learning system is given by Segal and Kephart (1999; 2000). They developed a Mail categorisation tool, *Mailcat* later to become *Swiftfile*, to aid users with filing activity. The tool predicted three folders in which it suggested users would like to file an email and provided short cuts to aid in doing so. To achieve this *Swiftfile* used a text classification algorithm comparing the content of each email received and those already filed (i.e. an IR based method). By

determining words which frequently occurred in one folder but infrequently in others, it was possible to distinguish between them. Looking for frequently occurring words within an email, it was then possible to determine which folder's profile was most represented. This principle of using words which occur frequently within a source but infrequently within a set is referred to as Term Frequency Inverse Document Frequency (TFIDF). An advantage of the system is that it automatically recommends where to file in accordance with users specific habits, and can adapt as they change. The system learns based upon a users existing filing structure, with folders naturally defined by the messages within them.

*Swiftfile* was integrated into Lotus Notes (IBM, 2008), with the principal of working in the background and offering the user a suggested file for each message, but allowing them to wave this. On a large sample of emails, although a small number of users, *Swiftfile* was tested to show around 80% accuracy to be achievable when allowed to predict 3 folder choices. Prior to *Swiftfile*, Boone (1998) presented and tested (in more limited settings) a conceptually similar tool *Re:Agent*.

Several tools have also been developed to use a Naïve Bayes (NB) classification model to file email (e.g. Rennie, 2000; Nenkova and Bagga, 2003). The principle of operation is that the probability of each word within an email then occurring within the different folders calculated. By combining probabilities for all of the words within an email, it is then possible to identify to which folders they are most likely to belong. Rennie (2000) chose to implement an NB classifier, because of the methods suitability to iterative learning. The tool was developed and reported to be in use by a community of around 40 members, and was shown to have a higher rate of accuracy than other tools, notably *Swiftfile*.

Koprinska et al., (2007) introduce a more recently evolved machine learning technique Random Forest (*RF*) to the classification of email. They reviewed, tested and evaluated RF along side existing tools such as

NB, support vector machines (SVM) and decision trees. Results showed RF to offer improved accuracy, learning and running speed.

The majority of the email classification models discussed here primarily use topics identified within emails. An exception to this is a tool developed by Cohen et al., (2004) who identify key verbs to do with activities such as confirm, propose and request, a speech act taxonomy is derived to achieve this. In an application using *TFIDF* and *SVM* they found that the detection of these 'speech acts' is more complex than topic identification but is possible. The implication is that this type of classification may be useful in aiding task management.

Crawford et al., (2002) tested a method to classify mail, using both rule based systems as well as *TFIDF*. The identity of the sender and the presence of keywords within emails were used in conjunction with a series of rules to determine likely folder choices. Again their tool *IEMS*, suggested rather than enforced folder choice, in much the same way as *Swiftfile*, as it was recognised this suited user behaviour and enforced the reading of email before filing. Further development of *IEMS* (McCreath and Kay, 2003) moved to using a purely rule based system for folder prediction using key word and sender information.

Rather than adopting the more common *TFIDF* and Bayesian approaches, Aery and Chakravarthy (2005) used a novel 'graph mining' based method to develop *eMailSift*. The principle was to identify interesting and repetitive patterns in the email body comparing incoming mail to that already in folders and proposing classifications based on this. An evaluation of the *eMailSift* was undertaken with various email sets to compare it with an NB classifier. It was shown that *eMailSift* was more effective particularly where there was a high degree of heterogeneity in the set.

The majority of the tools discussed have the ability to learn through their use, which has the effect of tailoring them to the specific user, this is

particularly beneficial as users' methods of dealing with and filing email have been shown to vary enormously.

Many of the classification tools have been developed widely but few, if any, have been developed and adopted. One explanation for this is that classifying email into folders is not an adequate or complete solution to effective retrieval. Where folders contain a large volume of emails there is still need for further sifting, often requiring keyword searching which in many cases is neither effective nor efficient. Most of the tools have been limited to putting emails into individual folders and nothing further. Using unique folders (rather than simultaneous locations) is still no more advanced than filing systems used for the last 100's of years and does not allow for more powerful combinational retrieval queries. Classification techniques also do not attempt to make content more accessible by changing the format of the information, such as by extracting content or adding contextual information to the emails. In many cases there is not even consideration to ensure that threaded conversations are filed consistently and together.

One suggestion is allowing emails to be classified with more than one label and thus enabling more powerful combined queries such as; every email sent by John, relating to project X and including minutes of a meeting. *Mailsleuth* (Cole, et al., 2003) used Formal Concept Analysis in applying a structured mark up to email texts to aid retrieval in this way. Their manual technique requires users to create and mark up concepts occurring within emails. A network of relationships between concepts and corresponding emails is consequently generated. This has the benefit of helping to retrieve mail through the use of combined concept searches but can also aid knowledge discovery throughout email corpora. Another example of the use of multiple folders to classify emails is given by Rohall et al., (2004) in the tool *Remail*, who also recognises the importance of maintaining threading.

Little work has been undertaken in the domain of email use specifically by engineers or designers, one exception is work by Kato et al., (2002). Using the IBIS principles (Kunz and Rittel, 1970), Kato et al., manually searched complete email threads, identifying where *issues* were raised and if a corresponding *decision* was made within the thread. Where this was the case they deduced the thread contained rationale. Their work did not go as far as identifying the positive and negative arguments and hence did not truly capture rationale.

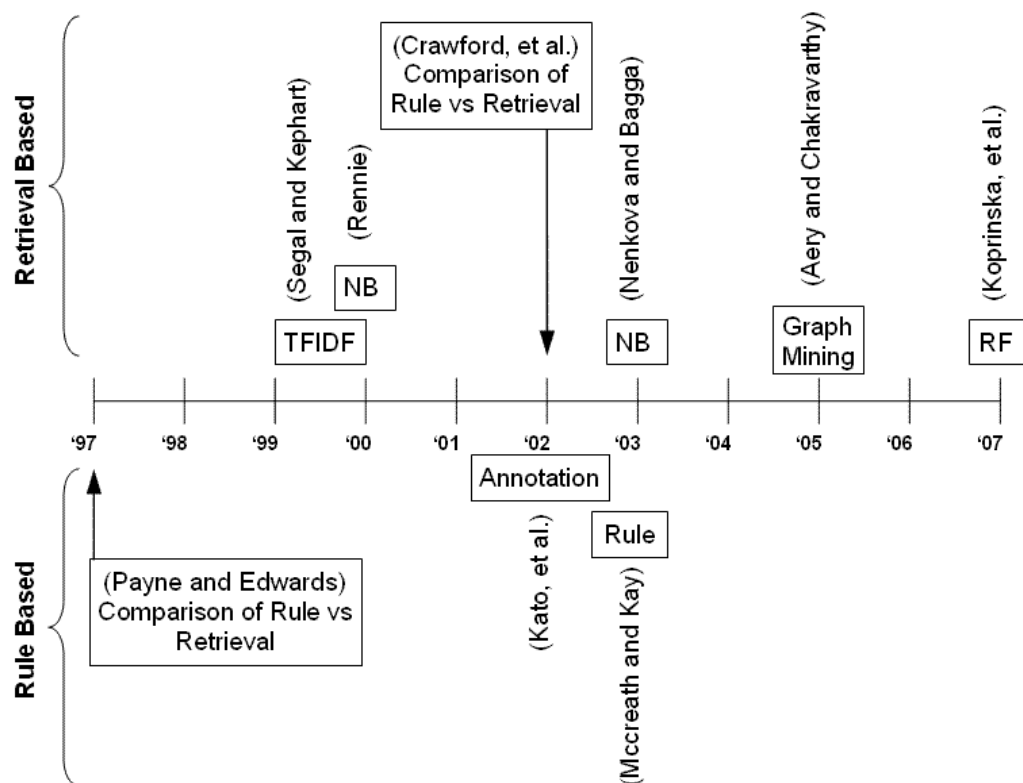
Despite this, the work remains useful in that it identified unresolved issues and unattended decisions in emails. Unfortunately the overhead on the part of users in implementing such a system is significant and most likely undesirable in a real life setting.

Minkov et al., (2006) present an algorithm for examining name disambiguation and email threading, which is aimed to be expanded at a linguistic level to improve IR (in email & documents).

Rather than classifying emails Schwartz (1999) attempted to hyperlink key terms or concepts within email text to an encyclopaedic style reference source or 'Organisational Memory' in the tool *Hypermail*. The system automatically proposed which words should be associated with known concepts in memory but similarly to other tools, allowing the user to support or modify these suggestions.

### **Summary of Classification Techniques**

This section of the review demonstrates the wide variety of techniques for email classification, using both machine learning (retrieval) and rule based classifications. Figure 3-2 depicts some of the specific works discussed on a time line.



**Figure 3-2: A timeline summarising some of the investigations into retrieval and rule based methods for email classification**

Both types of technique provide a viable option for classification however, the approach adopted by users still tends to be very one dimensional and has little difference conceptually to a paper filing cabinet.

Language processing techniques have improved and are capable of useful classification, as demonstrated by the tools discussed. However it is arguable that whilst computational methods cannot meaningfully understand text, the potential for these tools is limited.

### **3.1.4. Task Management**

Participants in a number of empirical studies have been shown to use their inbox as a place for keeping messages for the purpose of reminding themselves of tasks they must undertake. Despite the fact that email applications were not intended or optimised for this purpose, (Bellotti, et al., 2003) this behaviour has been consistently reported for some time (Mackay, 1988; Whittaker and Sidner, 1996; Gwizdka, 2002). A small



study exploring user requirements of email applications to aid task management was conducted by Bellotti et al., (2003). Seven principle needs were identified; these were the ability to:

- track multiple to do's,
- mark up the priority of tasks,
- manage threads over long time spans,
- manage deadlines and reminders,
- collate related items, switch easily between applications
- look through a task, rather than a folder oriented view.

The later requirement supports the view that email filing approaches (discussed previously) are not sufficient to deal with user needs in terms of task management.

Much of the work on task management has focused on how best to group and display information for users. For example Samiei et al., (2004) focused on improving task management by grouping of mail by threads, with the aim to present contextual and conversation history details in a highly visual manner in the tool *EzMail*. The suitability of user interfaces for task focused email was investigated by Gwizdka and Chignell (2004). They found that users' requirements were dependant on their ways of thinking such as having a visual memory or a working memory. This pointed to the need for a variety of interfaces for different users, a view that is recognised by others (e.g. Sow, et al., 2005).

A task-centric approach to an email client was proposed and developed by Bellotti et al., (2003; 2005) named *Taskmaster*. This boldly dispensed with the traditional inbox, drafts, sent mail, folders approach and instead grouped messages entirely by task. Threads of activity were used rather than straight forward threading of messages, and users could arrange tasks as sub tasks of another. The user was also granted the ability to put a meaningful name to tasks, rather than relying on subject headers. It

was also made possible to include supplementary sources of relevant information not originally sent by email, such as WebPages or documents. One of the benefits of the *Taskmaster* approach was that it became possible to generate task specific information, such as contact lists and summaries of actions to do. Additionally notifications of approaching deadlines and a visual approach to task requirements over time were possible. The testing of the tool was generally positive but the major problem was getting users to fully commit to changing from their established way of working and reliance on their existing email filing structure.

An alternative more holistic approach task management is taken by Kaptelinin (2003) presenting *UMEA* (User-Monitoring Environment for Activities). Rather than adding functionality to the email client Kaptelinin moves away to consider the whole working environment in an approach informed by activity theory. By monitoring user activity *UMEA* organises user resources including documents, folders and URL's from a project based perspective.

*SCOUT*, (Sow, et al., 2005) is unusual in that it has been developed to identify tasks within machine generated emails benefiting from their highly structured nature. A rule based approach is used to identify tasks implied by message content which are grouped contextually in terms of common business processes, such as account related. Although shown to be highly accurate, the utility of such an approach is clearly limited outside of highly structured machine generated emails.

Despite the work that has been undertaken in the area of improving task management within email, none of the tools presented appear to have been widely adopted.

### **3.1.5. Prioritising or ‘Triaging’ Emails**

Users who deal with large volumes of email have been seen to undertake a process of working through their inboxes, determining the most pertinent order in which to address emails and how to do so. In several cases the term “Triaging” has been used to describe this process (Venolia, et al., 2001; Bellotti, et al., 2005; Neustaedter, et al., 2005). The process of triaging undoubtedly relates to the classifying, task management and incoming email research presented so far; and could be considered at an intersection of these fields. It is worth noting that the automated filing of emails has received somewhat more attention than triaging although some of the processes undertaken are similar for the user. Notably triaging requires temporal consideration, with the ability to determine the order in which certain tasks must take place before others can occur, to fully automate such a process in free text emails would require a high level of computational understanding.

One technique has however been developed for the purpose of triaging emails, benefiting from specific and controlled circumstances and using classification methods. In order to address the issue of dealing with high volumes of emails to customer contact centres Nenkova and Bagga (2003) identified four types of email which were of interest. Firstly *Single* messages which did not require an immediate response, such as comments, suggestions for product improvements. Emails which did require a prompt response, such as queries were termed *roots*, subsequent emails to these exchanged between the contact centre and the customer were termed *inners* with the final email from a customer closing these conversations *leaves*. From the perspective of the contact centre it is clearly useful in being able to prioritise the order in which to deal with emails based upon these four divisions. The use of SVM and NB were compared in completing this task, with reasonably successful results. The relevance of this application to engineering organisations may increase as they move into the service sector, where they will more often manage their own customer contact networks. Although this tool

classifies email in much the same way as other tools discussed, its proposed application and the implication of identifying the four mail types, is such that this may be considered one of the only attempts to triage incoming email.

Whereas Nenkova and Bagga facilitated triage through classification Corston-Oliver et al., (2004) were concerned with task management. Their tool *Smartmail* automatically identified tasks in email and generated a reworded todo list for the user. This was achieved through the use of SVM's and Natural language Processing techniques (NLP). Although shown to have a reasonable level of precision-recall the techniques were not fail safe and risked tasks being missed or erroneously created. Despite the obvious utility of this tool, trials showed that users were uneasy about not reading the mail and conducting the process themselves.

It is the author's opinion that although research has identified the issue of "triaging," exploration of the processes involved and consequent proposals to aid it have not been made in the research field in general. Instead the use of email in task management, filing and incoming email have been given attention, which are all of relevance to effective email triage.

### **3.1.6. User Behaviour and the Effectiveness of Email**

To develop processes and tools for the purpose of improving email use, the understanding of user's interactions within the work place and email environment is seen to be critical. Several aspects of user behaviour relating to the processes of receiving email, filing, managing tasks and triaging have already been discussed in the previous sections. Other aspects of user behaviour which fit outside of these processes are now discussed; these relate particularly to the way in which email is used for communication and its implications socially, for knowledge sharing, and organisational management.

Much consideration has been given to the social implications of using email, rather than other communication media such as *face-to-face*, fax and telephone. The increased communication afforded by email has been recognised as positive in building relationships (O'Kane, et al., 2007) although it is acknowledged that it is not an ideal medium for introductions and initiating relationships (Wilson, 2002). The benefits teams get from email compared to *face-to-face* communication can be influenced greatly by the nature of the group and their willingness to interact. For most social processes email has been seen as being less effective than personal communication, however the intimacy rendered by emails is regarded as one of its strengths. One consequence of this is that individuals can say things to one another which they would not do in a *face-to-face* situation (O'Kane, et al., 2007) which although perceived as a negative factor may equally have positive consequences.

Email now plays a major role in intra-organisation communications and it is felt that as collections, messages contain a lot of detail about the context of the organisation and structures within it. Panteli (2002) explored this idea and found that email tends to reinforce existing power dynamics and hierarchies within groups; it should be stressed that the study was confined to academic institutions and conflicted with previous findings. An implication of email reinforcing power dynamics is to make processes less democratic, this has particular relevance for design groups using email who will need to debate, evaluate and make decisions. This is reinforced by Wilson's (2002) findings that email is neither good for negotiation nor reaching consensus. Despite these findings it is still recognised that email is used widely to support decision making and decision making processes (Lusk, 2006). It is clear from the literature that whilst email offers many benefits it should not be treated as an equivalent substitute for *face-to-face* communication. This extends to a wide variety of circumstances, particularly relevant processes for designers such as task execution, task selection, idea generation and decisions making have been specifically identified as less effective using email (Wilson, 2002).

In 1.4 the idea of information richness was introduced as: “the ability of information to change understanding” (Daft and Lengel, 1986). This is affected (in the context of face-to-face conversation) by factors such as tone of voice or facial expression which make the information ‘richer’. In the context of email Pantelli (2002) identifies that email richness is influenced through the way in which signatures, terminology and familiarities are used between parties; email text is more able to change understanding than documentation. The personal nature of the information increases its richness. Panteli (2002) also identifies that emails will provide information regarding organisational context; who talks to who and how they express their conversations. The premise that email contains such rich and useful information is fundamental to objectives of this thesis, such trying to improve accessibility and re-use of content. With the purpose of being objective two caveats to this are therefore made.

O’Kane et al., (2007) found that a principal purpose of emails was the transferring of information, passing views, reports or ideas from one person to another. The development of new understanding and creation of new ideas in emails was very much subservient to this. This does not exclude the possibility that valuable information is contained within emails, but does highlight the challenges faced in locating it. A secondary risk in regarding email as a rich media; is that its ease of use can have a detrimental on quality (Higa, et al., 2000). Circumstances surrounding how and when an email was sent, which are likely to be difficult to distinguish, may therefore affect the reusable value of content. It follows that consideration of these factors should be made when examining email records. Further to this it is acknowledged that such weaknesses could be overcome by better education of practices surrounding email user behaviour instilled by employers, improving effectiveness in managing processes and relationships (O’Kane and Hargie, 2007). The literature highlights that management have the greatest influence surrounding email policy and tool choice within their organisations, as well as how

much work is conducted with through email (Higa, et al., 2000); offering opportunity to affect improvements.

It is recognised that users are not generally keen to invest the extra time required in adding semantic value (through metadata) to their electronic media although this may have saved them effort later in time. McDowell et al., (2004) determined that although this may remain the case for most activities, where there is a large future benefit versus current cost, users are likely to be willing to use semantics. Three areas in which semantics could be used to improve email habitat were proposed to be: 1) *Process*, The handling of simple but tedious processes, such as automatic reply / invite to a function. 2) *Query*, answering other users simple questions, such as finding directions or a phone number. 3) *Updating*, automatically adding data to a webpage, database etc. from email content. .

### **3.2. Secondary use of email generated data**

In 3.1 the direct processes for which email is used were discussed. Email is however particularly suitable for deeper secondary exploration, as the content of communications and sender, recipient, time information are easily and automatically captured, unlike *face-to-face* meetings for example. Some ideas for exploiting email data in an automated way have hence been developed, either to serve a foreseen need, or because the possibility to do so is there. A number of these concepts and tools relate to how email data can be presented to show new or deeper meanings; in particular many of these adopt highly visual representations to achieve this, often termed visualisation techniques. One specific application of visualisation is the mapping of networks of communication within organisations. Along with analysis of such networks has been the quest to determine who within an organisation may have knowledge relevant to a problem. The following sections discuss firstly, general use of visualisation techniques and secondly how community structure and expertise location has been achieved.

### 3.2.1. Visual Representation of Email Data

It is often said that a picture is worth a thousand words; but can a picture be worth a thousand emails? As a way of condensing and representing patterns across corpora several researchers have developed 'visualisation techniques' to do just this.

One of the simpler examples of a visualisation tool was *Mailview* (Frau, et al., 2005) which was centred around message chronology, displaying as a point when each email had been sent on a time line. Users were able to change the view by filtering details such as senders or folder classifications. Whereas this tool was to assist with the retrieval of information another very different example aimed to monitor work flows over distributed locations. *Mapmail* (Nelson and Churchill, 2007) achieved this by displaying the sender, time and subject line of each email sent in a dialogue box on a map, enabling managers to observe patterns of communication.

From a more social standpoint Viegas et al., (2006) hypothesised "*that a visualization of email content constituted meaningful portraits of people's relationships.*" Demonstrating this, their tool *Themail*, presented the ability to view an individual's relationships and their content of communication with different people. This was achieved by allowing users to view keywords that appeared in their conversations, mapping them over a time line and enlarging them to emphasise their frequency.

This portrayal of information was termed the "haystack" and specific information sought within it, termed the "needle". As can be seen in Figure 3-3 this could highlight the passing of different events through related key words, or show with whom users were more frequently talking to about these subjects. It was viewed by Nelson and Churchill (2007) that although *Themail* allowed an overview to aid reading and browsing it lacked integration with the application view of an email client, an important feature to *Mapmail*.





several of these tools is of especial significance, as it could be relevant to understanding the changing states of email use through a design process.

### **3.2.2. *Community Structure and Expertise Location***

The concept of observing communities through electronic communication has already been mentioned in the context of Donath's (2002) visualisations of working groups, showing levels of involvement and dominance by players. Another permutation of community observation involves mapping the structure of communication patterns between individuals. The importance of so called 'social networking' to KM has long been recognised (Cross, et al., 2001) as it allows observation of how a given group of people create and share knowledge and consequently enables action to be taken upon this. Although the relationships that make up social networks can be formed in many different ways, analysis and mapping through email traffic has proved popular, probably due to the accessibility of data (particularly internal to companies). A further development of network mapping for KM is the location of subject experts, for which analysis of email content, more so than traffic has often been adopted. There has also been a wide uptake in social network mapping and expertise location for businesses through websites (e.g. [www.linkedin.com](http://www.linkedin.com)) and although these might provide an alternative to the email techniques now discussed, they demonstrate the value with which organisations regard such information.

The initial stage in examining a community structure is usually to map all the people as nodes (vertices) adding links (edges) between them where email communications have been sent. Links can be weighted in accordance with the volume (and possibly direction) of email traffic they represent; where frequencies are below a threshold, links are often removed to simplify analysis. The task of then identifying groups within the network is open to several different approaches.

One of the more established techniques uses the measure of *Betweenness Centrality*. Algorithms are used to determine the shortest possible routes between different nodes across a network; the more crucial a link or node is to these routes the higher its betweenness. By considering Inter-community edges as having high betweenness and Intra-community edges as having low betweenness it is then possible to break a network into sub groups. A good example of the use of this technique in the domain of email is given by Tyler et al., (2005). Using a dataset of one million emails and 400 users they were able to identify both formal and informal communities, as well as leadership figures.

Another technique which has been employed is *link-mining*. Nodes which have larger numbers of friends common to other nodes are considered to be more linked and hence central to groups. Rong et al., (2006) used a technique based on this to examine the community structure in the much publicised Enron Corpus (see Klimt and Yang, 2004).

Identifying and understanding social networks and communities of practice is seen as important in KM. Closely linked with this is the objective of determining who within an organisation or network knows what. Subject experts or general information are commonly sought, albeit subconsciously, using social networks; by asking questions of people and following referrals until a suitable person or answer is found. To supplement such processes, some organisations (particularly in knowledge work) create expertise databases. An interest has therefore arisen in ways of collecting information about expertise, particularly through automated examination of documents and email. Rationale for examining emails given by Campbell et al., (2003) is that they contain both network information about who people talk to, but also that people use them as a means of sharing knowledge or expertise.

Campbell et al., (2003) produced a model to identify expertise within an organisation through both the distribution pattern and content of emails. A principal difficulty in doing this is being able to distinguish between those

who know a lot about a topic and those who talk a lot about that topic. Using existing algorithms, content analysis looking for clusters of keywords identified the different topics emails related to. The topics discussed could then be associated with the different paths in a network map of the email traffic. Because people's knowledge of experts determines who they send email to, those receiving email are more likely to have greater expertise. An existing algorithm used to value Web pages, HITS, was applied to rate both the links in the network and also the authority of email text. This integrated approach of text and network analysis was applied to email corpora from two companies; although the utility of the technique was demonstrated it was recognised that further refinement was required to improve the quality of results. This tool would also benefit from ranking levels of expertise to improve functionality to the user (Bogers and Van Bosch, 2006).

Kim (2002) suggested that the use of keywords to identify topics for expertise identification in emails was not sufficient and instead applied *NLP* techniques to build user profiles or expertise maps. The model *EMNLP* (Expertise Model using *NLP*) evaluated information on both syntactic and semantic levels and was tasked to rank experts in order of their expertise based on a particular query – rather than by pre classified topics. Unlike Campbell et al., Kim did not consider the emails in the context of a social network, although this did not appear to diminish the system's ability to locate experts. It was envisaged that *EMNLP* could be embedded into the background of a user's working environment, encouraging its prompt and frequent use. Although more recent work has continued to address the area of expertise identification (e.g. Balog and Rijke, 2006) little further development has appeared.

The literature briefly described here has shown mapping social networks to be achievable using information generated by email traffic. It is also evident that there is room for further growth of these ideas, for example by measuring the strength of the links formed between individuals or sub-communities. It has also been shown that emails can be used to identify

experts in a subject area, to help for example with the fielding of questions. The ideas discussed in this section, and some of the previous visualisation ideas have been particularly relevant to KM. Although many of the ideas are distinct, it is envisioned that a more holistic approach integrating existing tools might produce a more useable output, which appears missing from the work to date.

### 3.3. Conclusions of Email Literature Survey

The two sections of this chapter discussed firstly the working processes of dealing with email and secondly the use of email generated data. In both sections aspects relating to research **Objectives 1** and **2** were considered. To conclude this chapter, as depicted in Figure 3-4 outcomes relating to these two objectives are discussed.

<b>3.1 Working Processes Dealing with Email</b>	<b>3.2 Secondary Use of Email Data</b>
<b>Objective 1 The Role Email Plays in Working Environments</b>	
<b>Objective 2 Tools to Improve Email Use</b>	

**Figure 3-4: The two sections of the review and the objectives now discussed in conclusion.**

#### 3.3.1. The Role Email Plays In Working Environments

Research **Objective 1** was: *To examine the existing research and determine what is known about the role that email plays both in engineering design and other working environments.*

None of the papers discussed has related specifically to the domain of engineering design, outlining the gap in existing knowledge which the research presented in this thesis aims to fill. Instead most research has related to generic business user groups and consequently the findings of

existing research presented here can only extend so far. With this in mind, findings of potential applicability and relevance to engineering design are highlighted here.

The most unequivocal finding of the review was that email is generally used to fulfil a large number of functions and has become a route through which many tasks in the modern workplace are undertaken; this serves to both email's appeal and down fall. Some of the most common activities include: file sending, scheduling meetings, customer contact and keeping people informed of progress. Because so many processes involve email, some users can be left overwhelmed by the volume of emails they consequently receive. Attached to this, is the finding that recipients of email tend to feel more pressured to respond than is necessary. This might apply particularly to engineering project managers who are likely to be recipients of many emails, particularly as their role requires them to be informed of many activities taking place.

It is also widely recognised that emails constitute a form of written record; as well as having data protection implications they may be of contractual significance. This again is applicable to engineering organisations, where ongoing negotiation of contract details is typical, particularly in larger projects as changes are made.

Although email is used for a wide variety of activities, some have been shown to be better suited to it than others, when compared to *face-to-face* communication. Furthermore the behaviour and attitudes of a group communicating by email can affect this. This suggests that caution should be taken in extrapolating findings from investigating email use by a design team. One of the processes that the literature suggests email is used for is to support decision making, although it may not be an ideal medium for this. The way in which people discuss, debate, evaluate, and choose may be significantly different from a *face-to-face* meeting. This might suggest that design teams could reach different outcomes depending upon their mode of working. The ease with which email is

used, or at least sent, has been seen to encourage communication to take place where it otherwise may not.

As argued extensively already, communication is fundamental in design activities. The question therefore raised is whether email encourages further and beneficial communication to take place, or whether it replaces other forms of communication where it is less effective to do so.

There is a certain amount of disagreement amongst the literature as to the potential value of emails information content. It is argued that people use email largely as a conduit and for sharing information rather than creating it. Along with this the ease of use is suggested to have a detrimental effect on quality. However, the conversational nature of emails means they provide a personally tailored presentation of information which is richer than that found in more formal reporting mechanisms such as documents. Emails are also considered to be a rich information source in the way that they include personalised expressions and explanations. The value of an email is relative to what is required from it, which may in part explain the discrepancies in views presented. It might then be fair to conclude that the intent and circumstances with which emails are created can be expected to affect their quality. Furthermore their value should be interpreted in the context within which they are proposed to be examined.

### **3.3.2. Tools to improve email use**

Research **Objective 2** was: *To examine the tools and techniques proposed to improve email use, including archiving and retrieval as well as general work processes.*

The review highlighted a range of tools proposed to aid existing methods of using email. Additionally to this some ideas for secondary processing of email data to extrapolate further benefits were introduced. All the tools were seen to utilise one or both of two distinct information types

associated with emails. Firstly the metadata or contextual information generated with each email, such as names of the sender and recipient and the time and date of sending; secondly the text, comprising the body or content of the email. The author acknowledges this as evidence to suggest that much useful information is associated and created with an email; furthermore this investigation should consider both the content and context within which emails are used to appreciate their true value.

The inbox has been described as the centre of the email client, with newly received messages and those already read but awaiting reply, further action, filing, or serving as reminders. The tools discussed have all worked within this existing framework and attempted to aid the user by adding automation to further processes or structuring the layout of the inbox. It has been generally recognised that tools should work in conjunction with the user, rather than carrying out steps such as filing without supervision.

A few approaches for task management have been identified and even fewer for triaging emails. It is felt that both would be highly applicable and beneficial in managing large engineering projects. The only tools to receive significant attention are those associated with the classification, filing and retrieval of emails.

The importance of IM in engineering design has been strongly argued in Chapters 1 & 2 and the management of email is proposed to be an integral part of this. Of the tools reviewed none were found to truly address email management *per se*, merely stand alone techniques to automatically classify for filing or to improve search and retrieval. Large engineering projects tend to use PDM systems (as discussed previously) which integrate information resources and contextualise them; research about encompassing email into such strategies was notably absent.

There has been a great deal of interest in the problem of automating classification, and the review has demonstrated a number of approaches,



the comparative evaluation of which is best left to the IR community and the fullness of time. Regardless of the optimal algorithm for classification, it is currently recognised that meaningful understanding of free text is far from computationally possible. More relevantly than reliability measures, is perhaps that classification is still regarded as sufficient for IR, and that few of the tools have attempted to add any additional semantic value to correspondence or allowed filing in multiple folders. Some of the visual techniques for reviewing email archives offer a novel alternative and it would be interesting to test and develop them further within a design setting.

A number of ways of using emails to derive community networks and locate experts have been explored. Based on the evidence of work with Converteam, it is believed that such techniques, which may broadly fit under the umbrella of KM, are of great use to engineering organisations. Where no record for decision making rationale exists, being able to locate someone with expertise in a particular system can prove invaluable. Tools for automated expertise location were reported to be of sufficient reliability that they may be of use, but still require much further development. A wider variety of methods were available for mapping networks, and also visualising communication patterns whilst relating them to subject matter. The main conclusion to be drawn was that the tools presented could be extremely valuable to engineering designers, but again better integration with other systems and more holistic approaches were needed to realise true benefits.

Of the many tools discussed the single biggest common link is the lack of an uptake into mainstream computer use, despite the fact that many have been researched by Microsoft and IBM. This can possibly be explained by the lag time between development of tools and their integration with finished packages. With a few highly dominant players in the computer software market, it is also unlikely that the ideas from small research groups will be taken forward independently as a realistic commercial venture. Attempting to market an email client, or an associated add on is

also likely to incur reluctance from IT managers, as any difficulties could easily cause large and costly implications to what is ironically a central system to operation.

In conclusion the review of tools for improving email use and reuse has highlighted the need for more holistic tools, which are better integrated not only within the email client, but also the general computing environment.

The perception from the literature is that users have a generally cluttered approach to their email use. Anecdotal evidence would suggest that this problem is as applicable to engineering designers as to any other working community. An opportunity therefore exists for the development and adoption of more integrated email clients, and to improve the behaviour of email users.

To enable this to be effectively achieved there is a need for a fundamental understanding of the role and use of email in engineering projects. The following chapters address this through the development of an approach for analysing emails and the application of this to case studies.

## 4.0 An Approach for Analysing Emails

Chapter 3 highlighted that there is little research which has been undertaken in terms of analysing the use of email in engineering design projects. **Objective 3** of this research was: *to develop and validate an approach for analysing the role and use of email within the context of engineering design projects and the overall design process*. This chapter presents the first phase of this objective by proposing an overall approach for analysing emails.

The premise for this research centres heavily around the need to gain a substantiated understanding, beyond anecdotal evidence, of how email is used in engineering projects. The need for such evidence is recognised in the computing research community, for example The Collaborative User Experience Group at IBM (Wattenberg, et al., 2005). Whilst most research to date surrounding email use has relied on interview and observation studies (O'Kane, et al., 2007) little has also considered the content of emails. In order to construct a suitable approach for a content based study of email use, this chapter reviews existing research, based on which an approach is developed.

### 4.1. Review of Existing Research Methods

An empirical study of how an engineering organisation uses email (including a correspondence content analysis) is a necessary and defining part of this research. Email dialogue presents an opportunity to explore both processes of communication and how information is used. A significant body of empirical design research has explored many facets of such topics. The following sections discuss the application of techniques including content analysis of transcripts and documents, diary studies and interviews for such research. Applicability to the investigation of email use is explored. The results of the review are used to inform the development of an approach for analysing email content.

#### **4.1.1. Studies of Design Communication**

Empirical studies exploring the process of design have been a major thread of research for some time (Cross, et al., 1992; Gero and Mc Neill, 1998; Stempfle and Badke-Schaub, 2002). A method which has been commonly utilised for this is protocol analysis, which elicits thinking processes from examining dialogue. With emails also representing a form of design communication, albeit asynchronous, such techniques are therefore of relevance.

A branch of content analysis, Protocol analysis is based on the well supported assertion that an individual's communication provides evidence of their cognitive reasoning (Ericsson and Simon, 1993). During a design scenario communication is recorded and transcribed. This script is then coded in accordance with a scheme, linking elements of text to key concepts. (Further discussion as to the development and validation of coding practices is given in Chapter 5). The concepts identified by the coding process then form the basis for further analysis. In Stempfle and Badke-Schaub (2002) for example, the frequencies of processes are explored as well as how they evolve within conversations, and across the entire design scenario. Such studies offer insight as to the processes and purposes of communication, but do not consider the details of the topics being discussed. Examples of processes might include analysis.

Some characteristics of the existing studies place limitations on their transferability to email corpus investigation. The settings for such studies (e.g. Stempfle and Badke-Schaub, 2002) have tended to be contrived design orientated situations, lasting a few hours and involving half a dozen participants, usually engineering students. In contrast industrial engineering projects often last for a number of months or years and involve the participation of many players including managers, administrators, technicians and marketers. It is argued that to properly understand how email is used in design projects analysis of real projects should be undertaken. Although parallels can be drawn between

contrived and real settings this study represents a significant change of scale. One implication is the difficulty in coding large volumes of text which is in any case a highly labour intensive process, particularly if the elements coded are small, such as clauses rather than paragraphs.

It should be noted that design communication has been studied in real and large scale settings. Rather than analysing the content of conversations, Eckert (2001) reports using interviews and observations to achieve this. This is a somewhat more practical approach than recording, transcribing, sampling and analysing conversation content in such settings. An advantage from the perspective of email analysis is that communication is recorded by default in a textual form.

#### ***4.1.2. Studies of Information Use***

Alongside the study of design communication (thinking), there has also been an interest in the use of information artefacts to support the design process (McMahon, et al., 2004; Aurisicchio, 2005). With email being a mechanism of information transfer and use such studies are relevant. Methods have included the use of content analysis (Lowe, 2002) as well as interviews, diary studies and observations (Aurisicchio, 2005).

Lowe (2002) classified phrases within documents as being technical terminology, contextual and product, resource or business related. The phrases collected within each of these groups were then assessed to explore what typified them. This approach was applied to engineering documents, such as reports, drawings, letters and emails. The emails examined were shown to contain lots of contextual information and company specific acronyms and tended to be written in an informal style. This method focused on distinguishing between the attributes of different information sources. A disadvantage of the method is that it requires coding on a very detailed level, consequently time consuming, and inappropriate for analysing a large corpus. Although not directly

transferable the method does however support the notion of using content analysis to examine emails and their use.

To learn how designers acquired information Aurisicchio (2005) asked engineers to compile a diary detailing each time they searched or requested information. This 'diary study' technique is particularly suited to collecting data in a workplace scenario over a long time period and from a number of participants. A primary disadvantage to applying this method to learn about how engineers use email, is the potentially high overhead placed on participants dealing with a large amount of email traffic. To obtain more detailed findings Aurisicchio also shadowed engineers for time periods of several hours each to observe their information using habits. Direction gained from prior study is felt to be important to make use of this technique. In a similar way to Lowe (2002) a coding scheme was used to surmise key characteristics, in this case from diaries and observations, rather than a document set.

Interviews were also used by Aurisicchio to gain a fuller understanding of how engineers operated, to contextualise and aid understanding of the collected data. They were also used to ask engineers for greater detail to help explain observed phenomena.

#### ***4.1.3. Study of Design Review Meetings***

The earlier discussion of design communication considered group working and informal meetings. Design review meetings are by nature more structured. Huet (2006) studied such meetings and coded transcripts using a predefined scheme. This highlighted, the type of communication (clarifying, debating), who was speaking, and which area of the design it impacted. Further details on the topic of discussion were annotated. This method captured the different elements of knowledge generated during design review meetings and was able to show how they evolved.

This study is particularly relevant to analysis of emails because it also investigates a by-product from an engineering process in a real life setting. The consideration of the topics discussed as well as how communication occurs is also favoured. Whereas design meetings consist of an extended period of synchronous communicative interaction, emails may represent only parts of a discussion, with other parts being discussed by other means.

#### **4.1.4. Conclusions**

A summary of the research methods discussed is presented in Table 4-1.

**Table 4-1: A summary of relevant methods for the analysis of email use discussed in section 4.1.**

<b>Method</b>	<b>Author(s)</b>	<b>Setting</b>
Protocol (content) analysis	Stempfle & Badke-Schaub (2002)	Processes and purposes of design communication in small groups.
Content Analysis	Lowe (2002)	Engineering document analysis e.g. reports, letters, emails, CAD,
	Huet (2006)	Design review meetings (transcripts of)
Interviews	Eckert (2001)	Communication in design groups
Diary Studies	Auricchio (2005)	Designers acquiring information
Observation (by shadowing)	Auricchio (2005)	Designers acquiring information
Statistics of email traffic	Nelson & Churchill (2007) Campbell et. al., 2003) and others in Chapter 3	Networks of people from email connections, volume and frequency of emails.

It is asserted that investigation of email use in engineering projects cannot be readily achieved in contrived settings, such as the ones reported. The variety of participants, locations, design complexity and timescales encountered in real projects cannot be easily represented. Although previous design studies in real project settings have used interviews and observations, the textual record created by emails provides an opportunity to explore them through content analysis.

To investigate both communication and information in engineering design a number of authors have analysed content using coding schemes.

These have usually been developed and agreed prior to their application to the document or transcript set. Coding approaches have identified concepts both relating to subject matter and processes (activities). The gathering of supplementary data has been undertaken in previous studies to aid with the further analysis of coding findings. Interviews, observations and diary studies have been used as such sources.

## **4.2. Research Approach**

Informed by the previous discussion of relevant research methods, this section presents the overall approach adopted for investigating email use in engineering design projects, depicted in Figure 4-1.

The adopted approach is centred on analysing the content of emails from a number of industrial engineering design projects. This was perceived to be an appropriate method given that emails are largely text based and that such techniques have also been demonstrated in engineering design studies of similar context (e.g. Lowe, 2002). A methodical approach to analysing content, provides an effective way of identifying consistencies, themes and patterns within collections of email messages (Patton, 2002).

In comparison with the other methods discussed in section 4.1, such as observation and diary study, content analysis of emails has the major advantage of being unobtrusive. By examining emails retrospectively it is possible to gain a direct insight into email usage, without risk of influencing user behaviour. It is also logistically simpler, enabling the study of large, complex, long-term engineering projects where real time observation would not be practical due to resource limitations.

It is important to recognise the limitations of content analysis. Although it allows a level of inference, such as to why an email was sent, it does not provide the full context to the sending or receiving of messages. Interactions outside of the email domain are not included. The time spent



with each email, effort committed and feelings of the user cannot be conveyed.

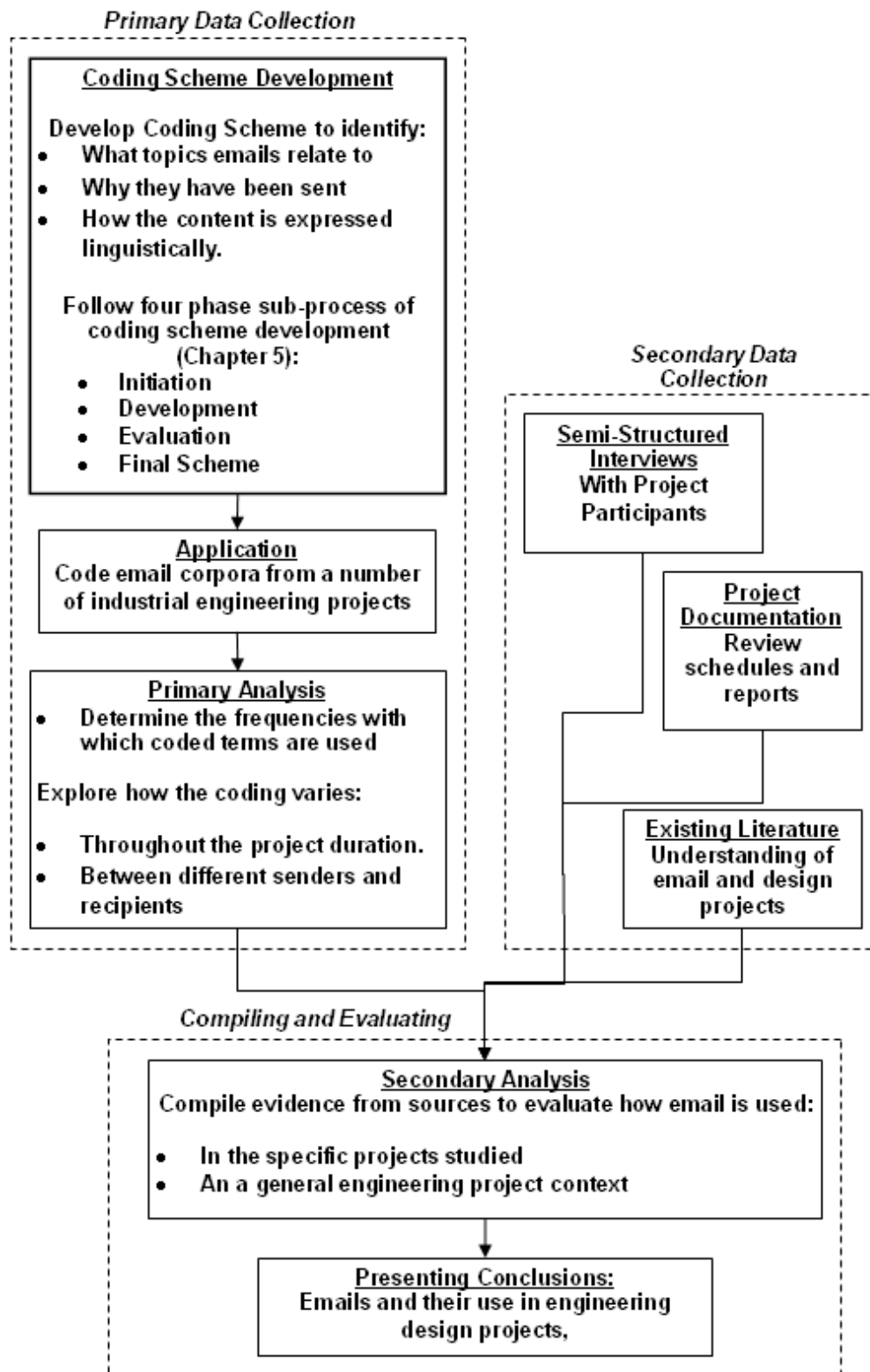


Figure 4-1: Research approach for the study of email use in engineering design projects

The overall approach followed, Figure 4-1, is now described in three key components. The first of these is termed the *primary data collection*, which involves the core content analysis. This includes the development of the coding scheme and its application to a sample of engineering project emails. The preliminary analysis, mainly presentation of statistical data, trends and initial observations, is conducted at this stage.

The next component is termed *secondary data collection*. This includes the gathering of evidence from engineers, project documents and existing research. This is used to contextualise the specific project(s) analysed as well as to provide more general background understanding.

The third and final stage is the *compiling and evaluating of evidence*. This uses the secondary data to further analyse the preliminary findings. The contextual background of the project is used to explain observed trends in the data. The relationships between results and findings of previous studies are also explored. Conclusions are then drawn in relation to the initial research questions.

#### **4.2.1. Primary Data Collection - Content**

The process of primary data collection has been divided into three stages. The first part of this is the development of a coding scheme which can be used to highlight the characteristics of engineering design emails. The second stage is the application of this scheme to a sample of emails from an industrial engineering project(s). The final stage is the primary analysis of the results; identifying how frequently terms occur and how this changes with time and between users.

#### **Coding Scheme Development**

A suitable and rigorous coding scheme is a critical part of the approach, the development of which Chapter 5 is devoted to. Prior to this development (and to enable its direction), the purpose of the scheme and the broad characteristics it needed to capture were identified.

In 4.1 it was shown that previous researchers had coded in relation to both the subject matter of text and the processes or activities inferred by it. As a progression of this three conceptual groupings were identified, *what*, *why* and *how*. The title of each of these groups is posed as a question, which the email coding scheme should answer.

***What topics does the email relate to?***

The subject matter of communications has been considered relevant. The topics of conversation have been annotated alongside existing coding (Huet, 2006). The domain of interest such as the product, resources or business has also been identified by coding (Lowe, 2002). It was identified that the variety of subjects discussed in emails is an important element in their portrayal.

***Why is the email being sent?***

A number of the coding schemes have been highly focused on processes, sometimes disregarding the subject matter. This includes thought processes (Stempfle and Badke-Schaub, 2002), reasons for information transactions (Auriscchio, 2005) and purposes for communicating (Huet, 2006). It is asserted that there is always a purpose for sending an email, and as such this should be identified.

***How is the content expressed?***

Having identified *why* information has been created, *what* that information is and, by virtue of an email's meta-data, *when* and by *whom* it was sent the remaining question posed begins *how*. A parallel is drawn here with Eckert (2001) who in exploring designer behaviour identified that it is important to understand how information is generated and created. This *how* is already partly known, the answer being via email, but the way in which this is written could reveal more. For this reason it is asked how is the content expressed?

## **Application**

Several of the methods previously discussed used contrived settings, often with student teams, for their studies (Valkenburg and Dorst, 1998; Stempfle and Badke-Schaub, 2002). These methods assume that findings can be translated to the real engineering environment, a reasonable assertion when examining specific interactions. However to investigate email use across design projects coding a corpus from a contrived exercise would not be a scalable approach. The design complexity, range of actors, expertise levels, parties, distributed locations and duration of real projects would be hard to replicate. For this reason the scheme developed for this research was applied to a number of engineering project email corpora obtained from industry. This was despite (and in acknowledgement of) the possible benefits of performing a contrived study where more holistic data collection and user participation can occur. Emails from design project corpora were sampled and coded. The sampling of every  $x^{\text{th}}$  email to represent trends amongst a larger corpus is an accepted technique (Krippendorff, 1980) and allowed a project with thousands of emails to be analysed with a realistic workload for coders. This approach has been applied in other studies where large collections of text have been analysed (Lightman, et al., 2007; Lianga, et al., 2008). In this study only single messages are coded and not email threads, this is discussed further in 5.6.1.

## **Primary Analysis**

The coded terms were presented with respect to a number of other factors. Similarly to Stempfle and Badke-Schaub (2002) the frequencies of coded terms were examined, along with how their use varies across the project duration. An exploration was also made as to how their use varied between participants with different roles in the project. When examining coding changes over time, Stempfle and Badke-Schaub considered 'macro and micro' perspectives. The macro-perspective in this case, how email use varied throughout a project, was most relevant to this investigation. The micro-perspective, how coding develops over a

conversation, was less of a priority than gaining the bigger picture, reflected in the coding scheme. This was particularly the case given that it is possible for only part of a communication exchange to be captured in an email, while the remainder occurs elsewhere, such as *face-to-face* dialogue.

#### **4.2.2. Secondary Data Collection - Context**

Although the approach was dominated by the collection of primary data through analysis of email content, the secondary data collection was of equal importance, with a two fold purpose. Firstly this contextualised the project; enabling a reasonable level of interpretation when coding emails. Required details included establishing the parties involved, objectives of the project and common acronyms. The second purpose was to enable further analysis of results, such as explaining observed phenomena. General perceptions of email use were also to be ascertained. This approach to support findings is similar to the one adopted by Aurisicchio (2005). The context gathering consisted of three aspects, interviews with project participants, review of project documentation and an understanding of email use presented in existing literature.

#### **Semi-Structured Interviews**

A semi structured interview uses prepared questions, to initiate open responses from interviewees. Spontaneous questions can be used to explore emerging topics further, in line with the interviews objectives (Wengraf, 2001). This method enables interviews to be structured enough to ensure that all necessary topics are covered; yet flexible enough to go into greater depth in areas most relevant to the interviewee. For this reason, semi-structured interviews were conducted with members of the project team. These included people from a variety of roles to gain as broad a perspective as possible. The questions explored an overview of the project, their role in it and their feelings and experiences relating to email use within the company. Full details are presented in Appendix A.

## **Project Documentation**

Project documentation was used to provide an overview of the project at different stages during its life. For this reason both superseded and final versions of documents were used. Of particular interest were project schedules which were used to provide an overview of key phases in the project. Minutes and reports from stage review meetings provided further detail. Project organisational charts were used to determine the links between different actors and parties in the project, and theoretical lines of reporting.

Exploring project documentation also enabled the reported events of a project to be compared with the actual events as described in emails. This gave a perspective of whether useful information is contained in emails which might otherwise be lost.

## **Existing Literature**

Chapter 2 presented an overview of design projects and particular issues pertaining to collaborative working and email use. This source provided a necessary background to understand the overall context of the project, but also to help explain findings.

Chapter 3 reviewed existing literature relating to email use, which was mainly reported in domains other than engineering. This evidence enabled the similarities between email use in engineering design and other contexts to be compared.

### ***4.2.3. Compiling and Evaluating Evidence***

Content analysis, interviews and project documents were correlated to evaluate the use of email within the specific projects (and company) studied. The final stage of the approach was the evaluation and presentation of findings with respect to email use more generally in engineering design projects. This was supported and compared with findings from the existing literature, presented in Chapter 3.

## **Secondary Analysis**

The secondary analysis explored two key areas. Preliminary findings were first interpreted in relation to the specific project(s) examined. This was used to explore the way in which emails represented characteristics of the project. It was aimed to explain why trends and patterns in email use occurred. The second area was intended to elicit findings from this, which are applicable to email use in engineering projects generally.

## **Presenting Conclusions**

The final part of the approach was the forming and presentation of conclusions, based upon all of the analysis undertaken. These were linked to the research questions posed in section 1.8.

### **4.3. Chapter Summary**

This chapter began to address **Objective 3** of the research: *To develop and validate an approach for analysing the role and use of email within the context of engineering design projects and the overall design process.*

To achieve this, a number of methods applied to the study of content analysis and in particular to design information and communication were discussed. These included the use of interviews, diary study, observation and content analysis via classification and coding. The findings of the review were that as part of the overall approach for analysing email use and content, the method of coding emails was most suitable for this research activity.

It was proposed that the coding scheme be centred on using content analysis to explore email use. This approach was selected due to emails having the availability of text for analysis and evidence of the use of such approaches and their success in previous engineering studies.

This scheme was required to identify three key features of the emails namely:

- 1) what topics were covered in the email,
- 2) why the email was sent, and
- 3) how the content of the email was expressed.

Figure 4-1 depicts the research approach which consisted of primary and secondary data collection, followed by compiling and evaluation.

Primary data collection was to involve the development of the coding scheme (Chapter 5) and its application to an engineering project email corpus. Characteristics of emails were to be coded allowing analysis of their frequency and how this changed with time, sender and recipient.

A phase of secondary data collection was to be used to draw further background evidence relating to the specific project and email use more generally. This included semi-structured interviews with project participants, review of project documentation and the use of existing literature (reviewed in Chapters 2 and 3).

Using these sources a secondary level analysis to explain findings in relation to the specific project and engineering design projects more generally was to be applied. Conclusions were to be presented in relation to the initial research questions.

Chapter 4 summarises the overall approach proposed for analysing the role and use of email. The following chapter explores the development of the coding scheme in detail to meet **Objective 3**.



## **5.0 Coding Scheme Development**

In the previous chapter the development of a coding scheme to mark up email content was identified as a key part of the approach to analyse emails. In this chapter the development of that coding scheme is presented.

Based on literature, some principles for the production, development, application and validation of coding schemes for content analysis are identified. Following this the method adopted in the development of the coding scheme is described and justified. The intermediate coding schemes and findings during this development then presented. A complete version of the final coding scheme is presented, including validation and exemplar findings. Finally, limitations of the coding approach in relation to attachments and threading are identified.

### **5.1. Principles for Scheme Production, Application and Validation**

This section discusses an array of the principles and guidelines for producing, developing, applying and validating coding schemes, as presented in the literature. Through these stages a number of multifaceted and interrelated factors are considered, and developing a coding scheme with consideration of them is as much of an iterative process as design.

To direct this discussion (influenced by Krippendorff, 1980) the scope and purpose of the coding scheme is first considered. The reliability of coding schemes and finally the validation and suitability of results is then discussed.

### 5.1.1. Scope

For the purpose of this work the term **scope** is used to describe the breadth and depth of the concepts that a coding scheme encompasses.

A key objective of coding (email) content is to explore the presence of recurring themes (Patton, 2002), which is achieved by describing characteristics of the text through a collection of labelled concepts. Labels which are either too broad or too narrow will reveal little. Codes may be developed in hierarchies or subsets to minimise such risk. Guba (1978) proposes the use of a convergent-divergent process. First; concepts which interrelated are identified and grouped. These categories are then expanded upon into further detail.

There is a possibility that a coding scheme may fail to capture relevant concepts which are present in the text. Assessing this is problematic for a number of reasons. The boundaries of a portion of text coded to a concept may be somewhat subjective, although necessary for context as words do not usually relate to concepts in their own right. Furthermore one portion of text might relate to more than one coded concept.

For example consider the sentences ***“Have you had any more thoughts about the fascia? I am considering changing the colour. Bob thinks it should be blue, this would improve legibility.”*** A hypothetical coding scheme could identify the concepts of ‘changing a design’ and ‘ideas’. Words relating to ‘ideas’ are *“thoughts,” “considering,” and “thinks”*. Key words relating to ‘changing the design’ are *“fascia,” “changing [the] colour,” “blue,”* and possibly *“improve legibility.”* If only these words were coded, phrases like *“have you had”* would be omitted but this would not necessarily identify a lack of capturing a relevant concept. There is an overlap here in the concepts and it would be very difficult to code each word of text to one or other of them. It follows that all three sentences could be coded by both concepts in their entirety. Although this would give an impression of completeness their may still be

a further outstanding concept, for example 'proposals' (suggesting the colour should be blue). The completeness of coding is therefore dependant on the scope and objective of the coding scheme.

Quantifying that a document is fully coded or that a coding scheme is complete is therefore impossible; much rests on the objective of the coding scheme. For this reason Patton (2002) proposes that researchers, observers and dataset providers should judge the scheme for its credibility, logicity and perceived inclusiveness.

Ultimately the scope of a coding scheme should be sufficient for the purpose of the proposed analysis. Lincoln and Guba (1985) upholds the virtues of adopting a guiding theory prior to scheme development. This ensures the correct questions are answered reasonably, rather than obtaining great detail which is not of interest.

### **5.1.2. Reliability**

Krippendorff (1980) suggests a **reliable** code will repeatedly produce the same results under the same conditions; regardless of how valid or useful these may be. Testing reliability requires some amount of duplication.

A facet of this is Inter-coder reliability, which measures agreement between two (or more) coders in their application of a coding scheme to a data set (Kurasaki, 2000). As well as demonstrating the general reliability of the scheme it allows more than one person to bear the burden of coding text, without unduly influencing results. It is acknowledged that the differing backgrounds, skill sets, training and perspectives of coders will influence their coding decisions (Patton, 2002).

Lombard et al., (2002) found that inter-coder reliability is widely underreported, and question the validity of any such work. They identify a number of different indexes for measuring inter-coder reliability, noting that simple percentage agreement does not account for coincidental

agreements. Acceptable values for agreement indexes are acknowledged to be subjective but it is suggested that indexes of 0.9 or higher are almost universally acceptable, above 0.8 suitable for most circumstances, and 0.7 sufficient for exploratory research.

To be reliable the application of codes must also be stable (Krippendorff, 1980). Researchers mark up strategies may evolve throughout the process and it is recognised that the point in time at which an article is coded may influence outcomes (Lincoln and Guba, 1985). Coding schemes should also be transferable to other document sets within the application domain.

### **5.1.3. Validity**

The **validity** of results is determined by their quality and also their representation of true facts (Krippendorff, 1980), a valid scheme also meets its purpose. As a qualitative approach the substantive significance of results are considered by way of a range of factors.

One of the key elements is comparison of the results. Findings which are consistent with existing knowledge, or match the expectations of the study are indicators of validity (Krippendorff, 1980). It is also desirable for the results to yield new depth to understanding and meet the purpose of the study (Patton, 2002).

Krippendorff also suggests that the coding scheme should be pragmatic, working under a variety of circumstances, such as with more than one email corpus in the case of this study. As part of this pragmatism the code should be logical and straightforward to apply.

A major influence on the validity of an entire coding study is the sampling of items for the set to code. The sample size must be sufficient to be representative of the entire case. To determine this, the 'split-half' technique is suggested; when a sample is split in half and results from

both parts match it is recognised to be sufficient. It is also important that the sample is evenly distributed over the dataset.

## **5.2. Method For Scheme Development**

The method by which the scheme was developed is depicted in Figure 5-1. The following sections describe this in four stages. The initiation stage grounded ideas and directed the scheme to answer the necessary questions. The development stage applied and tested these ideas before refining them. The evaluation stage verified the suitability of the developed scheme, before it was proposed in its final form. The issues of scope, reliability and validity were considered throughout.

### **5.2.1. Initiation**

As recommended by Lincoln and Guba (1985), the scheme was developed with a predetermined guiding theory to ensure it identified relevant characteristics so as to enable suitable analysis and findings. Three concept groupings were established, (in Chapter 4) which posed questions the coding scheme should answer. These were:

- What topics are discussed in emails?
- Why are they sent?
- How is their content expressed linguistically?

The method adopted treated each one of these facets individually. Terminology from existing coding schemes and taxonomies relating to each of the three questions was identified in engineering design literature. Based on these some initial structures for coding schemes were then proposed by the research team. This consisted of postgraduate students and academics with varying levels of industrial experience, all based at the University of Bath.

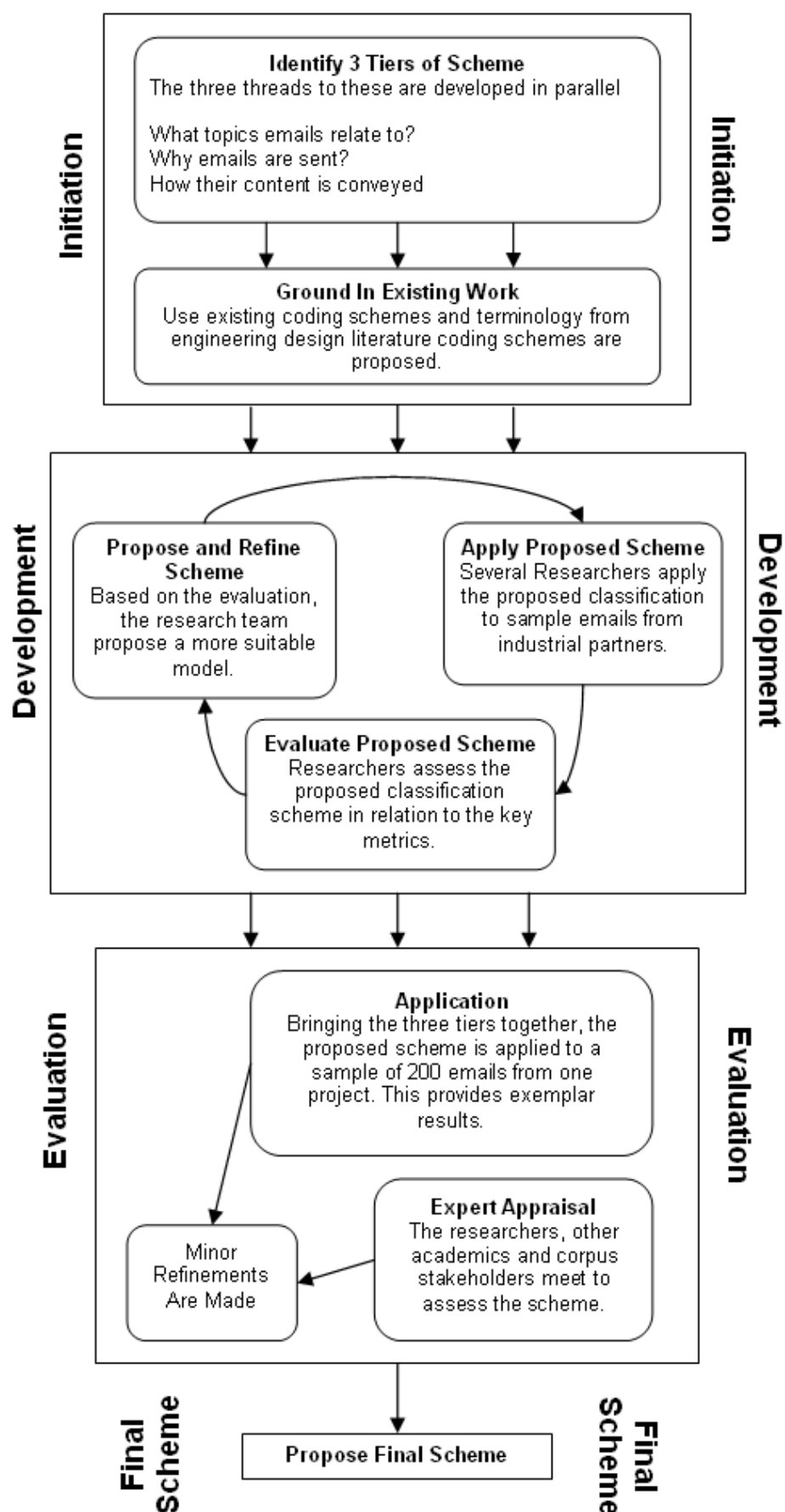


Figure 5-1: Process of coding scheme development, validation and finalisation.

### **5.2.2. Development**

These initial schemes were then developed further by a process of iterative refinement. Researchers applied the schemes both independently and jointly to sample emails. These came from a variety of engineering projects, to ensure the scheme was suitably generic for all engineering design.

Following the application the proposed schemes were evaluated in relation to the largely subjective criteria of their scope, reliability and validity, as discussed in 5.1. During discussion sessions the researchers compared their results and then proposed variations and refinements to the schemes. The amended schemes were then applied to further emails from the sample, and the process repeated until a consensus was reached that a suitable scheme had been developed.

Throughout this process the terminology from existing literature identified at the earlier stages was still used to ensure academic rigour in the grounding of the work. Where necessary, based on proposed scheme changes, further concepts from literature were sought.

### **5.2.3. Evaluation**

Once satisfactory schemes were developed for the elements of *what*, *why* and *how*, these were brought together to form an overall scheme. This scheme was then applied to emails from a single engineering project by three of the researchers. Exemplar results were then prepared showing frequencies of terms and how their use varied with time. A number of these emails were marked up by all three researchers allowing their inter-coder reliability to be calculated.

Patton (2002) suggested that three parties are best placed to judge the success of content analysis using a coding scheme. These are the developers, independent experts and someone with knowledge of the corpus. Hence the researchers, other academics with relevant expertise

and parties involved in the project detailed in the corpus attended a review meeting.

The issues of scope, reliability and validity were considered; as suggested in the literature (5.1). These were broken down, to a number of key criteria.

#### Scope

- Was the coverage broad enough and were any relevant characteristics failing to be captured?
- Was the level of detail sufficient, or should coding elements be broken down further?
- Were there terms which overlapped unnecessarily?

#### Reliability

- Was the reliability between coders sufficient? 0.7 was the suggested threshold.
- Was the scheme consistently applied, or did researchers coding behaviour change as time passed?

#### Validity

- Were the conceptual groupings identified in the coding scheme recognised to be 'truthful'?
- Was the scheme reasonably efficient and effective to apply?
- Were the results appropriate to answer the research questions?
- Were the trends reflective of expectations, based on previous studies and knowledge of the specific project?
- Were the trends shown found to offer new insight?

On the basis of these criteria and the corresponding review meeting, final minor adjustments were made to the scheme. If significant these would have required a repetition of this stage of the process.



#### **5.2.4. Final Scheme**

To conclude the process of scheme development, a final version of the scheme was presented. This included descriptions and definitions for each term and how it was to be applied. An overall guide to the coding scheme was also presented. To support these, examples of the coding applied to real emails were produced.

### **5.3. Detailing of Scheme Development**

The coding scheme was developed in accordance with the method previously described. This section summarises the key developments of that process and aims to present background and reasoning behind the final scheme. The three aspects of the coding scheme are presented in turn; firstly *what* topics emails discuss, secondly *why* they are sent and finally *how* their content is expressed. For each of these the grounding literature used to influence the terminology and structure is discussed and its preliminary application described. Findings and reasoning for key developments from this initial state are then presented. To avoid repetition, and ambiguity arriving from changes during development, the definitions for terminology within the coding scheme are included only once with the final version of the scheme at the end of this chapter (see page 128).

#### **5.3.1. What - Development**

This portion of the coding scheme was intended to answer the question: *What topics do emails discuss?* An initial structure was based predominantly on work by Huet (2006) and Lowe (2002). Their schemes were appealing since although coming from two different and relevant perspectives their terminology largely overlap. Huet's scheme was designed to capture dialogue, which it is argued emails represent a form of (Kim, 2002), whereas Lowe's scheme has been applied to engineering documentation and correspondence wider than solely design.

The first terms used similarly by both Huet and Lowe are *product* and *process*, to describe the artefact and act of designing it respectively. Notably these terms are also used by other authors (e.g. Ahmed and Wallace, 2003). Huet and Lowe also used the term *resources*, encompassing physical, human or financial aspects.

Only one disparity existed, where Huet refers to *external* factors which influence the design, Lowe refers to *business issues*. It was felt that these are similar, and bear an overlap with one another other. This collection of terms, brought together in Table 5-1, was applied independently and cooperatively to a sample of emails by three researchers, before findings were discussed.

**Table 5-1: Initial coding scheme for What topics were discussed.**

Product	Process	Resources	External	Business
The artefact	The design process	Resources used in process	Influences on process	Business issues (possibly affecting process)

## Findings

It was found that the term *product* was clear to apply and occurred in a significant proportion of emails. As a point of definition it was noted that the product may be non-physical, such as a piece of software.

The term *process* was heavily used, and was felt to cover too broad a range of concepts to differentiate between. This arose because the design process was not explicitly mentioned and did not form a topic of discussion in its own right. It was found that any verbs occurring in the text could be interpreted by the coders as processes because they referred to actions. It was proposed that, for the purpose of analysing emails, it would be clearer to identify *project* related activity. The term

*project* reflects the domain in which an action takes place and is less ambiguous and more easily defined than a process.

The external and business terms were shown to be similar, and it was observed they reflected influences by the *project* on the design, and from the *company* on the design. A small proportion of emails were found which discuss *resources* explicitly. It was viewed however that these would be better considered as a facet of the *project* or *company* rather than a group in their own right.

### **Further Development**

Based on these findings it was proposed that the terms *process*, *resources*, *external* and *business* be replaced with the terms *project* and *company*. Identifying that an email is discussing the project, rather than the artefact which it produces was felt to be intuitive. Similarly company related dialogue, abstract from a specific project, is conceptually distinct.

A second tier of coding was proposed, such that *product*, *project* and *company*, formed three categories, within which a set of further terms existed. To select and develop these terms a further search of existing literature was used.

Definitions and concepts surrounding the *product* have been thoroughly researched and described within engineering design. It was hence possible to select a suitable set of pre-existing terms as a basis on which to develop. Li and Ramani's (2007) taxonomy was selected for this purpose.

Similarly structured or grouped terms were not found directly for the *project* or *company* facets. Using the texts discussed in the review of engineering projects and organisations in Chapter 2, (Field and Keller, 1998; Gray and Larson, 2000; Haberberg and Rieple, 2001; Smith, 2007) suitable terms were identified and grouped.

The new scheme, an early version of which is shown in Table 5-2, was then incrementally developed through application to emails by the researchers, as previously described.

**Table 5-2: Development of coding scheme structure for what topics were discussed.**

Product	Project	Company
Function	Team	Stakeholders
Material	Risk	Financial Resources
Environment	Cost	Tools / Methods
Performance	Schedules	Human Resources
Manufacturing	Contracts	Physical Resources
	Deliverables	Knowledge Resources
	Time	Practices & Procedures

## Findings

The hierarchical approach enabled top level categories of *product*, *project* and *company* to be better defined by the features that comprise them, improving coder reliability. It was also appreciated that there is a trade off between the reliability of a scheme and detail it goes into. Having two levels made this compromise easier, allowing one level to be more reliable and concise, the other to explore greater detail.

This layout and terms were found to be more intuitive and distinct to apply, improving reliability. It was also felt that the findings were more easily interpreted. The terms within each of the categories were developed, particularly influenced by different characteristics of different projects whose emails were sampled. Note again that the final version of the scheme, incorporating aspects of *what*, *why* and *how* is presented at the end of the chapter.

**Table 5-3: Terminology identifying purposes for which communications occur.**

Ref.	Term Grouping	Ref.	Term Grouping
1,2,5	Informing, Sending Information Receiving Information	2,3 4, 5	Managing, Management, Control, Planning Coordination
8	Requesting Information	8	Justification
6	Reflecting	1, 4	Analysis
1,2,3 4,7	Evaluating	2,4,7 8	Decision Making, Decision
1,4,7 8	Solving, Solution Generation	2,5	Problem Solving, Resolving Problems
4, 7	Goals	7, 8	Constraints, Constraint Negotiation
1	Confirmation	1	Comparison
8	Conflict Resolution	8	Negotiating Clarification
[1]	(Auriscchio, 2005)	[2]	(Huet, 2006)
[3]	(Sim and Duffy, 2003)	[4]	(Stempfle and Badke-Schaub, 2002)
[5]	(Lusk, 2006)	[6]	(Valkenburg and Dorst, 1998)
[7]	(Pahl and Beitz, 1996)	[8]	(Eckert and Stacey, 2001)

**5.3.2. Why – Development**

This portion of the coding scheme was intended to answer the question: *why are emails sent?* As a first step towards identifying purposes of sending messages, terminology was compiled from both empirical and theoretical studies of design communication, Table 5-3. Schemes which had been intended for both formal interactions, such as meetings (Huet, 2006) and informal interactions, normal working (Auriscchio, 2005) were reviewed. The breadth of these ranged from design specific orientation (Valkenburg and Dorst, 1998) to general engineering practices (Sim and Duffy, 2003). This array of terms was then applied to a sample of emails, as per the method, and findings discussed between the researchers.

## Findings

The collection of terms was found to be suitably diverse, capturing a variety of communication purposes, without omitting any. It was however felt that the process of applying the terms was not particularly pragmatic, given their variety and a lack of defined relation between them. As a result the principle outcome of the application, was that redundant terms should be eliminated and that the remaining terms should be presented within more structured groups. It was envisaged this would also make the scheme more reliable to use.

It was observed that a number of the communication purposes related to *transactions*; these could be further divided into two groups. *Information* handling related transactions encompass requests for information, sending and receiving information. *Management* related transactions direct or request other people to take action of some form, beyond a purely information handling process.

The observation of *information* transactions was less than straightforward, given the incomplete representation emails give. For example it was difficult to ascertain whether information in an email had been previously requested. It might have appeared in another part of the thread which was not detailed, or have happened in a casual conversation.

The remainder of communicative purposes were found to relate to *problem solving behaviour*, such as generating solution ideas and giving evaluations. Given the significance of problem solving in design activity (e.g. Simon, 1969; Goel, 1995) and the ambiguity in existing literature as to how well email might support this (compare Wilson, 2002; with Lusk, 2006), such a grouping enables valuable insight.

The two categories proposed *Information and Management Transactions* and *Problem Solving Behaviour* may be further supported by the work of

Medland (1992). He identifies four types of communication relating to: delegation (*cf.* Management Transactions) awareness and reporting (*cf.* Information Transactions) and problem handling (*cf.* Problem Solving Behaviour).

### Further Development

Based on the proposal, terms were grouped into transactions, with the subsets of *information* and *management*, and *problem solving* behaviour. The version presented, Table 5-4, reflects this and shows a reduction through redundant terms compared with the starting point (Table 5-3). Acknowledging that it may not be possible to determine whether information has been freely initiated, or sent in response to a request, the term *informing* encompasses both. Facets of *management* transactions were found to provide an unnecessary level of detail and complications. Especially given that the activities involved, such as planning, were captured by the *what* part of the scheme.

**Table 5-4: Three proposed groupings of communicative purposes**

Transactions		Problem Solving Behaviour
Informing	Information Transactions	Exploring
Clarifying		Debating
Confirming		Identifying Goals
Requesting Information	Management Transactions	Identifying Constraints
Managing		Solutions
Including control, planning, directing		Evaluating
		Decision Making
		Reflecting

## Findings

The scheme presented allowed details of interest to be captured, as well as presenting more general characteristics through the use of groupings. Management and information distribution have been identified as key uses of email in generic contexts (Wilson, 2002) as has problem solving to design activity. The scheme therefore presents a suitable framework to evaluate previous research. Developments to the scheme mainly centred around definition of terms, particularly because the concepts are quite abstracted from the text.

For completeness it should be noted that more fundamental analysis methods, such as Speech Act Theory (Bach and Harnish, 1979) were considered. Although there is the potential to provide more detailed analysis, considerable time, effort and training are required on the part of the researcher. It was not felt that the added benefit was a worthwhile compromise for either initial research into email use or contributing a useable scheme for engineering research.

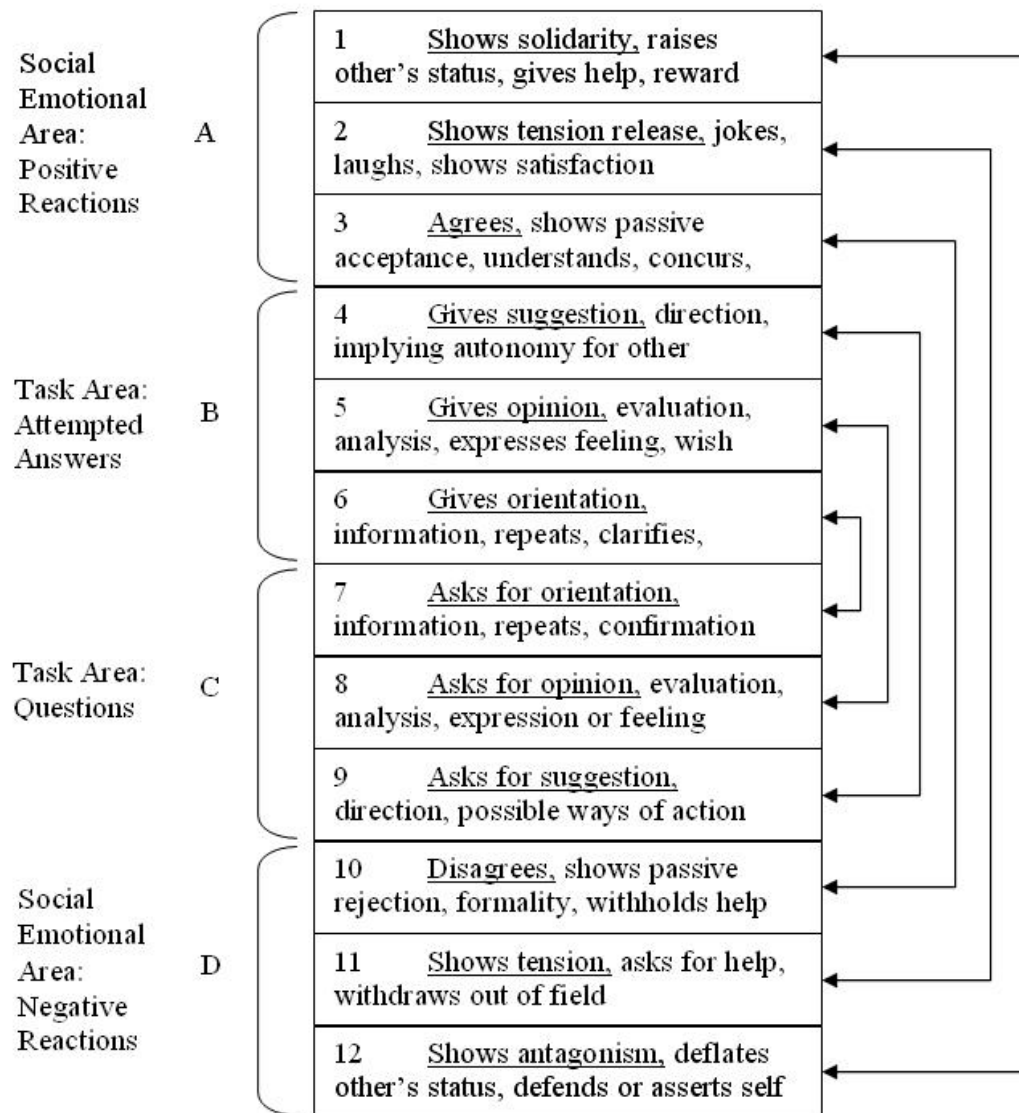
### **5.3.3. How – Development**

This portion of the coding scheme was intended to address the question: *how is the content of emails expressed?* With the intention of finding if anything could be learned from the language used and towards the characteristics of discussion.

A number of established approaches were explored, the foremost of which was the process for interaction analysis developed by Bales (1950; 1951). Bales used content analysis (via coding) to explore communication within small working groups. This method showed how types of phrase changed as discussions evolved towards decisions. It also showed how certain players dominated the group. Although Bales developed this technique to analyse co-located groups, it has been used to examine computer mediated communication between virtual groups (Hiltz, et al., 1980; Reid, et al., 1996) and indeed to compare this with *face-to-face*



working. In the engineering design domain Bale's approach has been favourably reported in comparison with other methods for examining meeting communication (Gorse and Emmitt, 2003) and has been subsequently applied (Gorse and Emmitt, 2007).



**Figure 5-2: Bales categories for analysing interaction processes, adapted from Bales (1950)**

Bales coding scheme, Figure 5-2, captures both social and task related characteristics. Both of these facets are particularly relevant to studies of designers, who work together solving problem based tasks. Three types of positive socio-emotional reactions are described, (group A) along with their negative equivalents (group D). Three types of question and

matching responses (group C and B) describe information relating to the task.

A number of other techniques to explore how communication is expressed were considered. Amongst these, Huet (2006) applied an approach labelling each sentence as a question, answer, statement, or feeling. It was proposed that a collection of such elements would comprise communication purposes (*why*), such as clarifying. This mark up has the advantage of being straightforward to apply.

Also considered were latent semantic techniques. Previous application of these (Dong and Vande Moere, 2006) has been used to identify how congruence develops throughout discussions and also how a groups' knowledge converges. Although the use of such techniques can offer detailed insight, their application is non trivial.

## **Findings**

Having applied the various techniques Bales' method was found to be the best suited and sufficient without modification. The combination of social and task analysis of Bales' method, along with its previous application to computer communication and design environments gave support to its adoption. It was also found to be no more time consuming to apply than other methods, whilst providing a rich level of detail.

It was noted that the difference between *what* and *why* parts of the scheme was clear, but the distinction between *why* and *how* was less obvious. The *how* content is expressed, coded using Bales' approach should always analyse discussion at the clause level, abstract from the overall purpose or purposes of the message. These should generally be identified on a paragraph basis as a *why*, although some clause length *whys* will still exist, particularly elements such as decisions. It was recognised that the purposes of emails (*why*) might be constructed

through components of *how* they were expressed; however it was relevant or beneficial to formally and theoretically link these.

## **5.4. Evaluation of the Scheme**

The final part of coding scheme development, as described in the method, was the combining of the *what*, *why* and *how* sub-schemes. These were then applied to a corpus to obtain exemplar findings and demonstrate the suitability of the scheme. As suggested by Patton (2002) the scheme developers, corpus provider and academic peers then judged the final output. To achieve this a review meeting was held, attended by five academics from a number of universities along with representatives of the company from which the corpus was obtained. The researchers who had developed and coded emails using the scheme presented the results. All then participated in a critical discussion. As an output of this minor developments to the scheme, particularly definitions, were made. These are reflected in the final version of the scheme presented at the end of the chapter. Findings of the evaluation are now presented in relation to the questions regarding scope, reliability and validity criteria posed in method, section 5.2.

### **5.4.1. Scope**

*Was the coverage broad enough and were any relevant characteristics failing to be captured?*

The use of the three aspects of *what*, *why* and *how* was felt to give broad and suitable coverage. No missing themes were identified either during the final application or by the review panel. The capture of interactions through *why* and *how* was viewed to be a major asset of the scheme.

*Was the level of detail sufficient, or should coding elements be broken down further?*

The two tier approach of the *What* part of the scheme was felt to be beneficial in capturing finer detail and similarly for the groupings used with the *why* part of the scheme. It was found that although the scheme had sufficient detail to annotate specific portions of text, the view of what the whole email contained was most appropriate

*Were there terms which overlapped unnecessarily?*

There was initially a considerable overlap between the *product*, *project* and *company* categories. Although this was noted to be useful, because emails could relate to more than one of these, it was felt that there was still too high an overlap caused by ambiguity. Improvements to definitions and clarification of terms within each category, were implemented to resolve this.

#### **5.4.2. Reliability**

*Was the reliability between coders sufficient? 0.7 was the suggested threshold.*

The figures for inter-coder reliability were 0.75 for 'what' subject matter the emails related to and 0.65 for 'why' the emails were sent. This was considered to be acceptable giving that the coders had a low level of training and experience and that further refinements to the scheme were anticipated. It was also noted that the proportional use of each term was very similar between researchers.

*Was the scheme consistently applied, or did researchers coding behaviour change as time passed?*

Coders noted that they felt that their use of the coding scheme could vary throughout application.

To mitigate against both of these factors (inter-coder reliability and consistency) it was suggested that in a large scale application, coders

should regularly re-train by coding emails together. It was also suggested that emails should not be marked up in chronological order, as such bias might affect observations of how email use changes with time.

#### **5.4.3. Validity**

*Were the conceptual groupings identified in the coding scheme recognised to be 'truthful'?*

The grounding of terminology and structure within existing literature assured the conceptual rigour of the scheme. Peers agreed that the concepts which the scheme captured were well founded.

*Was the scheme reasonably efficient and effective to apply?*

The scheme was generally pragmatic and experienced coders were able to code an average email in 10 minutes. The most time consuming process was found to be coding the elements of *how* content was expressed, due to its mark up at the sentence level. This facet of the scheme was also observed to be more valuable for exploring the micro rather than macro-perspective of email use. For this reason it was proposed that in later case studies a large number of emails should be coded with the *what* and *why* to obtain an overall picture, and a smaller subset also with *how* to learn about the intricacies of communication.

*Were the trends reflective of expectations, based on previous studies and knowledge of the specific project?*

The findings generally aligned with the company's expectations and knowledge of the project. For example a significant proportion of emails discussed *company* type information, unsurprising as this was an in-house project. A large proportion of emails related to sharing information, and to management, as reported by (Wilson, 2002).

*Were the trends shown found to offer new insight?*

A key insight was that a low proportion of emails showed *problem solving* and of these, very few showed any decision making. This confronted an

assumption by the company that important decisions were being recorded in emails, and raised the possibility that rationale is being lost.

*Were the results appropriate to answer the research questions?*

It was shown that emails relating to the *product*, peaked toward the end of the project schedule. This and similar views demonstrated the possibility of using emails to monitor the progress of projects, and that as set out in **Objective 4**, change in email use over time can be captured.

## **5.5. Final Scheme**

This section presents the final version of the coding scheme and a guide to its application, including examples. The scheme is discussed in the order in which its three aspects, *what*, *why* and *how*, are applied.

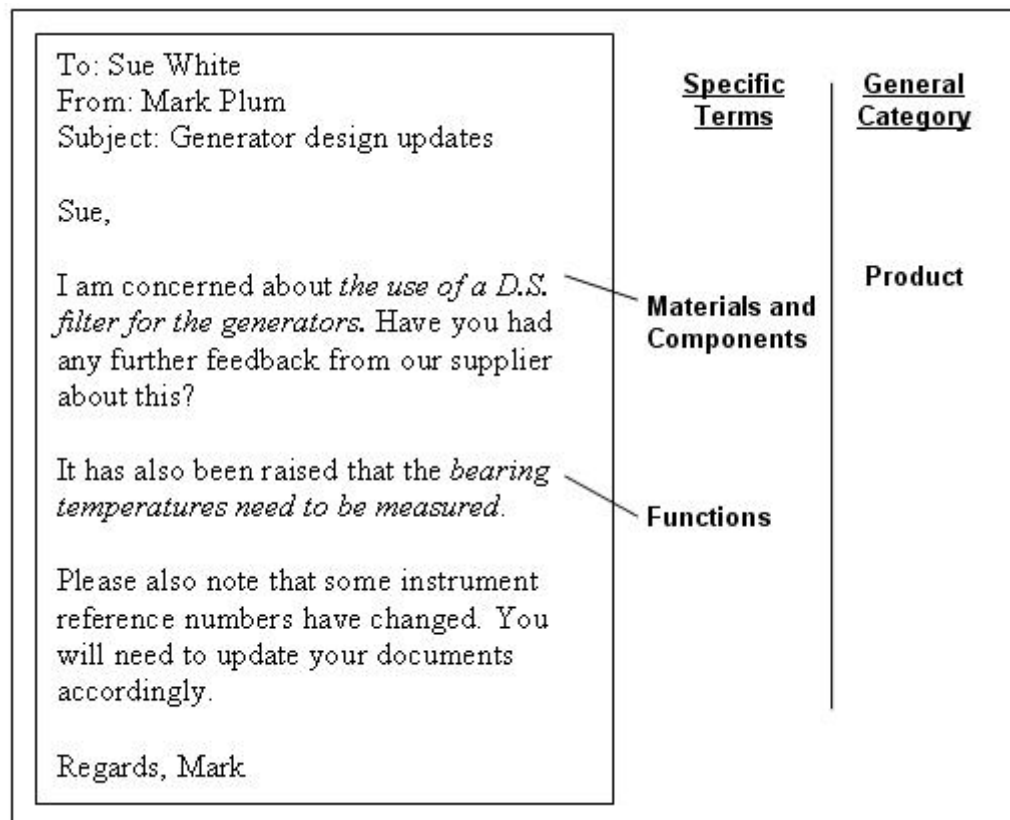
Throughout the coding process, the coder should bear in mind any contextual information which can be reasonably ascertained from the document. This might include the text in the subject field and anything known about the sender and recipient(s) i.e. is the email inter, or intra-company. Coders should make reasonable inferences from what is presented but not assumptions which are unfounded.

### **5.5.1. Final Scheme: What topics do emails discuss?**

The first stage is the coding of *What topics* emails discuss, the scheme for which is presented in Table 5-5. Having read the email, the coder should, based on the predominant themes of the message, determine whether it generally relates most to the *product*, *project* or *company*. If it significantly relates to more than one category, then more than one code may be allocated. Following this the coder should identify further details from within the selected category. In the example below Figure 5-3 an email related mainly to the product discusses *features* and *materials*.

**Table 5-5: Coding terminology and definitions identifying what topics email discuss.**

What topics does the email discuss? Subject matter it relates to		
<b>Product</b> The output of the project; it may be a physical artefact or software.	<b>Project</b> the domain within which the product is created.	<b>Company</b> the sponsors or facilitators of a project.
<b>Functions:</b> Things the product must do. e.g. Be fast	<b>Risk:</b> Assessing likelihood and weighting implications	<b>Economic Issues:</b> Costs and efficiency, market and product selection
<b>Performance:</b> How well the product achieves its functions.	<b>Plans:</b> Management of phases, activities and tasks,	<b>Human Resources:</b> People, availability, allocation, training, replacing
<b>Feature:</b> The quality or characteristic with which the function is achieved.	<b>Team:</b> Team selection, development	<b>Physical Resources:</b> Ranging from offices to equipment
<b>Operating Environment:</b> Objects that interact with the product	<b>Quality Management:</b> Quality, standard or expectations	<b>Financial Resources:</b> Cash, Assets, Borrowing.
<b>Materials and Components:</b> Materials and component selection and characteristics	<b>Cost:</b> Financial arrangements at the level of the project, rather than specific component costs.	<b>Knowledge Resources:</b> Current ability and stored information.
<b>Manufacturing:</b> Consideration of manufacturing, assembly and transport.	<b>Time:</b> Durations or deadlines. Any link or reference to time.	<b>Tools and Methods:</b> Specific testing and modelling techniques.
<b>Cost:</b> Consideration of costs particularly unit costs.	<b>Manufacture:</b> Arranging manufacture, planning manufacture, in the context of the project	<b>Practices &amp; Procedures:</b> Accumulated by the company, often developed through experience
<b>Specification:</b> Formal requirements definition for the product/design. Or requirements for sub / super components of the product.	<b>Contracts:</b> Legal arrangements involving two or more parties setting out what is required from the project, often specifying costs and time.	<b>Stake Holders:</b> Such as: share holders, customers, directors and their culture and politics. They must have more than an interest in the company but in the project.
<b>Ergonomics:</b> User Interaction with product	<b>Delivery:</b> The delivery or provision of a specific component or sub-system.	
	<b>Milestones &amp; Deliverables:</b> Targets to be achieved, or which have been achieved, related to formal stages within the project.	
	<b>Documentation or Knowledge Resources:</b> Reference to general documentation resource, most likely PROMIS or mention of Knowledge Management process specific to the project.	
	<b>Administration:</b> General administration related to the project, but not distinctly captured by one of the other terms above.	



**Figure 5-3: Coding What topics the email discusses: the general theme is discussion of the product, and terms highlighted within this category are components and functions.**

### 5.5.2. Final Scheme: Why are emails sent?

Applying the *what* aspect of the scheme first gives the coder a chance to reflect on the overall nature of the message. The next stage, marking up the *why*, requires a deeper level of interpretation. Table 5-6 presents the coding scheme; terms are listed as *problem solving contributions*, *information transactions* or *management transactions*. The term *contribution* is used in relation to problem solving to emphasise that even a small part of a problem solving process should be identified as an occurrence of problem solving in emails.

The coder is required to first identify which bottom level terms are demonstrated within the email; such as *exploring* or *clarifying*. This is opposed to identifying the top level categories and then their facets, as was undertaken marking up the *what* aspect of the scheme. Further



processing can then surmise which types of transaction are demonstrated, (*information* or *management*) and whether there is evidence of a *problem solving contribution*.

**Table 5-6: Coding terminology and definitions identifying why emails are sent, with respect to transactions or contribution.**

Why are the emails sent? Their purpose	
Problem Solving Contribution	Information Transaction
A contribution to a problem solving process. More active than simply passing information.	A straightforward transfer of information or something enabling the transfer of information.
<b>Goal Setting:</b> Identifying where the design is, and where it needs progressing to.	<b>Informing:</b> Sharing, presenting or distribution information with others. No response is required. It is passive.
<b>Constraining:</b> Imposing boundaries with requirements and desirables	<b>Requesting Information:</b> Direct request to another party to provide information, or further information. Including explicit responses to requests for information.
<b>Exploring:</b> Discussing possibilities and ideas, invoking suggestions. A return is expected from the recipient.	<b>Clarifying:</b> Clearing up misunderstandings (both requesting and in response). Asking for explanations, resolving a general lack of clarity.
<b>Developing Solutions:</b> It may encompass one or more of the following stages: searching, gathering, creating and developing solutions. Presentation of solutions for comment is also encompassed.	Management Transaction
<b>Evaluating:</b> Judging the quality, value and importance of something.	An effort to achieve action by asking or telling for things to be done.
<b>Decision Making:</b> Considering key factors from evaluation and possible compromises to form decision.	<b>Managing:</b> Includes arranging, directing and instructing. Implies action (such as a response) needs to be taken. Including process management outside of the organisation, e.g. prompting arrangements / meetings with third parties.
<b>Reflecting:</b> Reflecting upon a design/product decision or process already adopted or occurred. Reflecting may question whether a new or further problem now exists.	<b>Confirming:</b> Confirming or requesting confirmation of something
<b>Debating:</b> Discussing opposite views.	

Coders should highlight an appropriate portion of text for the coding term they are allocating. This may be a paragraph, or a few words alone. In particular, references to problem solving may only be evidenced by a short phrase. The quantity of words tagged by a particular label does not necessarily represent their significance, and assumptions should not be drawn on this basis (Lincoln and Guba, 1985). If this were the case coders would have to give much consideration as to where to start and finish their selection of text, overlapping concepts also present an issue.

Instead the coding determines which terms do and do not occur within the email. Highlighting specific passages aids the coding process and enables a return to the message later for further review. Figure 5-4 demonstrates the coding of purposes in a portion of email text. The first paragraph shows a confirmation, but a decision made as part of this is simultaneously coded.

	<u><b>Specific Terms</b></u>	<u><b>General Category</b></u>
<p>To: John Green Cc: Oliver Mustard From: Ruth Scarlet Subject: ---</p> <p>John,</p> <p>Following our discussion last week I can confirm that we have selected, and are ready to proceed, with ordering the roller bearings.</p> <p>Could you please liaise with Oliver in purchasing to make sure this is undertaken by the end of the week.</p> <p>Thanks, Ruth</p>	<p><b>Decision Making</b></p> <p><b>Confirming</b></p> <p><b>Managing</b></p>	<p><b>Management Transaction &amp; Problem Solving Contribution</b></p> <p><b>Management Transaction</b></p>

**Figure 5-4: The coding of why emails are sent: the purposes of confirming and managing are highlighted.**

### 5.5.3. Final Scheme: How is email content expressed?

The final element of the scheme is coding how content is expressed, through the application of Bales' (1950) method. Table 5-7 presents the twelve possible tags which can be applied to each clause or sentence; these are grouped into socio emotional and task categories, with all terms having a polar opposite. The scheme is applied at the term level by the coder, with the groupings used to aid navigation of the scheme and in later analysis.

**Table 5-7: Coding terminology and definitions used to identify characteristics of how email dialogue is expressed.**

How is email content expressed?			
Socio Emotional Terms		Task Related Terms	
Positive Reactions	Negative Reactions	Sharing	Requesting
<i>Shows Solidarity</i> raises other's status, gives help, reward	<i>Shows Antagonism</i> deflates other's status, asserts or defends self	<i>Gives Opinion</i> , evaluation, analysis, expresses feeling or wish	<i>Asks for Opinion</i> , evaluation, analysis, expression or feeling
<i>Shows Tension Release</i> jokes, laughs, shows satisfaction	<i>Shows Tension</i> asks for help, withdraws out of field	<i>Gives Suggestion</i> direction, implying autonomy for other	<i>Asks for Suggestion</i> direction, possible ways of action.
<i>Agrees</i> , shows passive acceptance, understands, concurs, complies	<i>Disagrees</i> , shows passive rejection, formality, withholds resources	<i>Gives Orientation</i> , information, repeats, clarifies.	<i>Asks for Orientation</i> information, repetition and confirmation

Examples in Figure 5-5 show both social responses, such as *showing antagonism* and task responses such as *giving orientation*. These labels are applied to sentences (presented as paragraphs for clarity in the figure).

<p>To: Jan Blew From: Steve Black Subject: Bearing Inspection</p> <p>Jan,</p> <p>The motor bearing inspection on the first unit is expected to be required on the 5<sup>th</sup> June. This is based on testing phases being completed in time.</p> <p>Would you please check for the availability of someone and let me know if there is a preferred date and time for this,</p> <p>Regards,</p> <p>Steve</p>	<p><b><u>Terms</u></b></p> <p>} <b>Gives Orientation</b></p> <p>} <b>Gives Suggestion</b></p>
<p>To: Ruth Scarlet From: Mark Plumb Subject: Re: problem with hinges</p> <p>Regrettably this was a totally foreseeable issue, which I raised several times in the early stages of the design!</p> <p>Given the circumstances however I feel your solution is most prudent.</p> <p>How likely do you think it is that we can have this resolved before the next review meeting?</p> <p>Mark</p>	<p>} <b>Shows Antagonism</b></p> <p>} <b>Agrees</b></p> <p>} <b>Asks Opinion</b></p>

**Figure 5-5: The coding of how email content is expressed using terminology from Bales.**

## **5.6. Limitations of Coding Approach**

The coding approach is limited in that threaded emails and attachments are not directly addressed. These falls outside of the scope of this thesis and were not properly supported by the datasets used for analysis in the following chapters. This section therefore discusses the issues of threading and attachments, outlining the extent of the limitations and providing thoughts for further work.

### **5.6.1. Threaded Emails**

This section begins by differentiating between threaded emails and replied using quoted text. Following this a discussion is made around the lack of threading in the case studies and the limitations for analysis.

#### **The difference between email threads and quoted text replies.**

A *thread* is a collection of emails in reply to one another. These are normally linked by grouping emails of a common subject heading. An example of a threaded email client is Google™ Mail. Figure 5-6 shows a view of a threaded Google™ Mail conversation with the subject line “Your Invoice/Booking Confirmation for 0403.” Threads may include forwarded emails and the inclusion of several parties. In the example, the original email and subject line has come from an airline, whilst the following conversation is sent between two friends planning a trip.

Not all email clients support threading. Instead many will quote the original email being replied to below the new response. This is shown in Figure 5-7, with a view from the University of Bath’s webmail client. As is common, chevron (>) symbols are used to denote the quoted text.



Figure 5-6: A threaded email conversation viewed in Google™ Mail

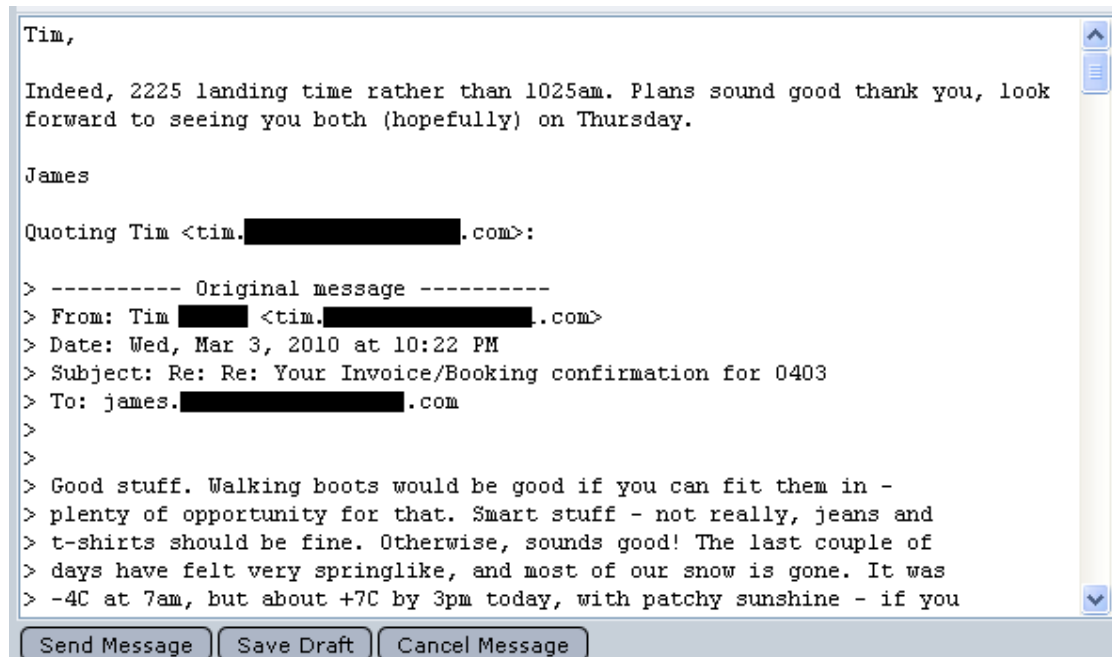


Figure 5-7: A quoted text email reply shown in the University of Bath's webmail client.

### **Limitations of coding single emails**

As with a number of email clients, the databases of the industrial partner did not maintain the threading between emails, which is largely based on similarity of subject headings. This was further complicated by the different systems used by external partners to the projects. As a result of this there was no threading in the email collections investigated in the pilot or major case studies (Chapters 6 and 7). In many emails however there was quoted text.

Many emails were effectively stored twice in the databases as they were repeated in the form of quoted text in other emails. To avoid duplicating coding and distorting results a strategy was devised. Only the principal text in each email was coded. Any quoted text would only be used to provide context to inform the coder as to the overall gist of the principal message. This is discussed further in 6.3.5.

The principal limitation in not coding emails as complete threads, is that no exploration can be made as to how discussions develop and evolve over the course of a complete email conversation.

It was already identified in 4.2.1 however that this is not the objective of the coding scheme. The focus of this thesis is to investigate the use of email at the project rather than conversational level. For this reason the coding scheme identifies the macro-perspective, how coding evolves over the project, rather than the micro-perspective, of coding changes over a conversation. The coding of single emails (using any quoted text to provide context) is supported by this approach.

This undoubtedly leaves a gap for future studies to explore coding at a conversational level, using complete threads of messages. For such an investigation the use of an argumentation or IBIS type approach to coding, as discussed in 2.1.2, may be more appropriate than the broader purposes for email sending presented in the existing coding scheme.

The coding of single emails does not have a detrimental affect on the results produced by the coding scheme. The approach identifies how frequently, when in the project and by whom certain types of email are sent. This can only be achieved by the sampling of single emails (as in the case studies) or by the coding of every email in corpus.

#### ***5.6.2. Limitations of not coding attachments***

Attachments are files which are sent from one person to another alongside an email. The principal purpose of the message may be to share the file by sending an attachment, or the attachment may provide relevant information to be considered alongside the message.

Approximately 21% of emails in the corpus originally contained attachments. Again, due to limitations of the databases, the attached files or links to them were not consistently maintained or available. As a result it was decided not to code individual attachments.

When attachments are sent, email is being used as a file sharing mechanism. In such cases, the purpose for sending the email (as described in the coding scheme) would normally be informing, i.e. sharing information. Hence, although the attachments were not coded, the purpose of them as alluded to in the emails was identified. In this sense, attachments have been considered by the coding approach.

It still remains however that a limitation of this thesis is that the detail of attached email content was not investigated. To do so would have produced inconsistent results as a comprehensive data set of attachments was not available. Future work would benefit from exploring in more detail the use and content of attachments.



## 5.7. Chapter Summary

Chapters 4 and 5 addressed **Objective 3** of the research: *To develop and validate an approach for analysing the role and use of email within the context of engineering design projects and the overall design process.*

General principles for developing and evaluating coding schemes were presented, based on literature of qualitative research methods. Considerations were framed around the key aspects of scope, reliability and validity.

The method for developing the scheme was then described. This included the use of terminology from successful approaches (Lowe, 2002; Huet, 2006) to form a basis for the scheme. This was then iteratively developed through application of the coding to sample and review emails by a number of researchers. An application of the finished scheme to 200 emails from a project, and peer review meeting ended this process.

A summary of the scheme's development using this method was then described. This followed separately the *what*, *why* and *how* aspects of the scheme. The validity of the final scheme was discussed in relation to the scope, reliability and validity metrics outlined earlier.

The full and final version of the scheme was presented, along with guidelines for its use and examples. (Further examples are given in Appendix B). To conclude, limitations of the coding approach in relation to threading and attachments were identified.

## **6.0 Pilot Case Study: Approach, Validation and Preliminary Findings**

A pilot case study was undertaken, in anticipation of the major case study, reported in Chapter 7.

This primarily served to empirically validate and evaluate the overall approach, as well as reaffirming the suitability of the coding scheme. A refined approach could then be used for the major case study. Exemplar findings would also serve to guide the way in which results were analysed in the major case study.

By examining a project which had significantly different characteristics from the major case study it was also possible to establish the applicability of findings, beyond the specific project and to engineering design projects in general.

The project background for this pilot study is now presented, followed by details of how the approach was applied. Results are firstly presented and discussed in relation to validating the approach. Key results and general findings are then highlighted.

### **6.1. Project Context**

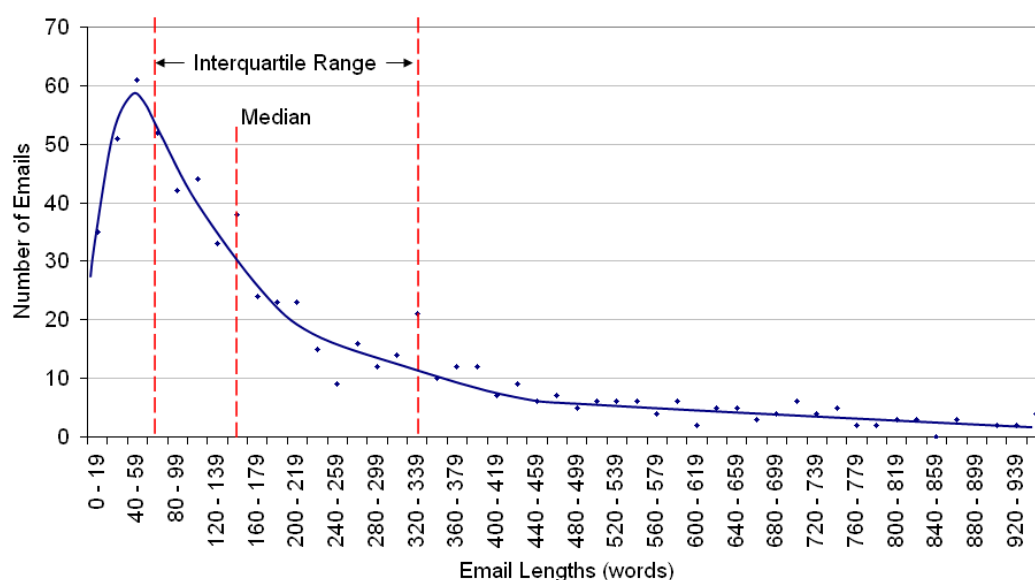
The project related to the development of a software product, used in the process of systems engineering by a large company. This, a new version of the software, was developed in-house to be used extensively across the organisation. Although such software development is an ongoing process, the core activity of this project took place over a two year period.

Although a large number of people in the company were consulted and involved during the project, the main project team consisted of fifteen members. These included eight software engineers, under the direction of

two lead engineers, with one project manager overseeing the process. The activity was conducted over two sites, one in the UK and one in Asia.

The email corpus created by the project contained around 700 emails, collected over the two year period. These messages were sent by 63 people to 158 recipients.

A distribution of the lengths of these emails is presented in Figure 6-1. The median length of an email was 140 – 159 words. A quarter of emails were relatively short (less than 60 words) and half of emails were between 60 and 339 words long. A total of three quarters of emails were between 0 and 339 words long.



**Figure 6-1: Distribution of email lengths in the Pilot Case Study corpus**

## 6.2. Application

To reduce the corpus to a more manageable number of emails for analysis a sample of 225 from the original 700 was taken, which was then coded by three researchers. In order to be representative this was distributed evenly across the corpus, with coders marking up alternating emails. Twenty of the emails were coded by all three researchers, in order to measure their inter-coder reliability.

These emails were fully coded using the *What*, *Why* and *How* elements of the coding scheme presented in Table 5-5, Table 5-6 and Table 5-7. As discussed in the coding guide, 5.5, the *What* and *Why* elements were coded to appropriate sized units of text, most frequently paragraphs. Owing to the flexible approach allowing units of any size to be coded, certain *Why* elements, such as requests for information or confirmations were often only sentence or clause length. The *How* element of the coding scheme applied terms to each sentence within the email.

The coding was undertaken using the software package NVivo (QSR, 2008), a qualitative research tool. This allowed any sized portion of text within each source (in this case an email) to be allocated to a given code. It was then possible to identify to which emails, and hence how many, each code had been applied. The coding process took around ten minutes per email. Given the time taken to setup and prepare the data sets, requirements and overheads, the whole process took around three days for each coder.

### **6.3. Discussion of Validation**

The pilot case study provided an opportunity to validate the overall approach for analysing emails, as well as the specifics of the coding scheme. The validation was important because changes fed into the major case study.

The metrics for assessing coding schemes, previously discussed in Chapter 5 are reviewed in turn: inter-coder reliability, stability and significance. The logistical aspects of the coding process are also referred to. A brief evaluation as to the value of the results is then made, followed by identifying the limitations of the approach.

### 6.3.1. Inter-Coder Reliability

The concept of inter-coder reliability was introduced in Chapter 5 as a method for measuring the consistency between coders. In this case study the three coders each coded the same 20 emails as a sample to test. The inter-coder reliabilities were then calculated using Cohen's Kappa (Cohen, 1960), and average values for the three coder comparisons are shown in Table 6-1.

**Table 6-1: Inter-coder reliability of top level codes  
calculated using Cohen's kappa.**

<b>Code</b>	<b>Coefficient</b>
Product	0.71
Project	0.71
Company	0.73
Information Transactions	0.66
Management Transactions	0.63
Problem Solving Contributions	0.61

The inter-coder reliability was found to be just sufficient (around 0.7) for the top level codes, although it was recognised that improvement was needed.

The instances where coders had disagreed were reviewed and discussed by the three coders. In most cases it was found that the coders had identified a further valid topic with which the others agreed with but had not identified initially. It was not often that two coders disagreed outright with each others coding choice. The varied experiences and backgrounds of coders is accepted to be a reason for their identifying different concepts (Guba, 1978).

It was observed that across their own wider samples, the three coders applied each label in a similar proportion of cases. This showed that the scheme was fundamentally sound, and it was judged that with further

training coders would achieve a greater degree of consistency in the major case study.

To mitigate against inter-coder discrepancies in the major case study, it was maintained that researchers should code alternate emails. This is especially important for the tests which examine how email use changes over time.

### **6.3.2. *Stability of coding***

Coders reported that they felt their use of the coding terms changed as they became more familiar with using them. There could be a tendency to fall into the habit of using certain terms. It was suggested that coders should initially code a number of emails from a project, not to be used in the analysis, to help train and familiarise them with the coding scheme and data set. It was also suggested to cycle through the sample, rather than coding all of the emails from the beginning of the project on the first day, and end of the project two weeks later.

### **6.3.3. *Significance and Logistics***

The pilot case study gave an opportunity to trial the logistical aspects of coding and to consider the statistical significance of results, level of detail they explore and the time taken to practically achieve this.

Approximately every third email in the corpus was coded, giving a sample which was sufficiently representative of the corpus. It was recognised that fewer emails could have been sampled to achieve a similar result. This would reduce the time taken to analyse a larger corpus.

The emails took an average of ten minutes each to mark up. A great deal of this was due to the coding of the *How* element, which required each individual sentence to be tagged, whereas other labels tended to cover paragraphs. The level of detail provided by coding in this way was high

but it was judged this could be reasonably compromised in order to make coding a quicker process for the purpose of dealing with a larger corpus.

#### **6.3.4. *Value of the type of results generated***

The first set of results concerned the frequency with which each of the codes was applied to the email sample. These results were judged to provide a strong representation of the purposes and subject matter of emails in the corpus. The top level of classifications, reported with the most confidence, were of most value.

The results which showed how email use varied from person to person, were of little value in this case study, as there was little variance from person to person. However it was viewed that in a project with more complex dynamics and roles this would change; as such this type of analysis was judged to be worthwhile undertaking for the major case study.

The final element of the results explored how email use changed over time. In this case study there was evidence that the role of email changed throughout the project. Identifying this was judged to be of value in understanding email use, particularly where further secondary data could be used to explain observed trends.

#### **6.3.5. *Limitations***

As with a number of email clients, the company's database did not maintain the threading between emails, which is largely based on similarity of subject headings. This meant that when coding, emails were treated as stand alone objects. It was recognised that the previous and subsequent emails in a thread add context which aids understanding of the message and aids in coding. It is therefore a limitation of the technique that sometimes emails will not have been as well understood

and therefore not as well coded as they may have been if full email threads were always available to examine.

A number of the emails also contained attachments. The method via which the email corpus had been stored meant that attachments were not easily accessible to analyse. In some cases the attachments will undoubtedly have contained important information or provided context to aid understanding of emails for coding. Although this is a limitation of the approach, it would not have been practicable to properly investigate attached files. It was judged this was less important than coding a range of emails.

The level of secondary data collected for this pilot study was limited. Informal communications with the project manager offered significant further understanding into how the project and company operated. As such, a thorough exploration as to the value of comparing results with secondary data was not possible for this case study.

## **6.4. Results**

This section presents the results generated in the pilot case study. The primary purpose of these was to aid in the evaluation of the overall approach. These results also demonstrated the potential for analyses made possible by the coding process.

The secondary purpose of presenting these results was to enable some basic findings to be drawn; which could then be used in the further analysis of the major case study.

The frequency of coding application, its variation between users and changes over the project duration are now presented.



#### 6.4.1. Frequency of Coding Elements

The first set of results presents the frequency with which terms in the coding scheme were applied to emails. These are presented as a percentage of the sample, and should hence be a representative percentage of the analysed corpus as a whole. It should be restated that because emails are normally coded to more than one concept, the sum of percentages in some figures will exceed 100.

The results detailing *what* topics the emails discussed are first presented, followed by the purposes for *why* they were sent. Finally *how* this content was expressed is detailed.

#### Topics Which Emails Contained

Figure 6-2 depicts the main topics groups *Product*, *Project* and *Company* which emails relate to; and shows the overlap of emails which relate to more than one of these.

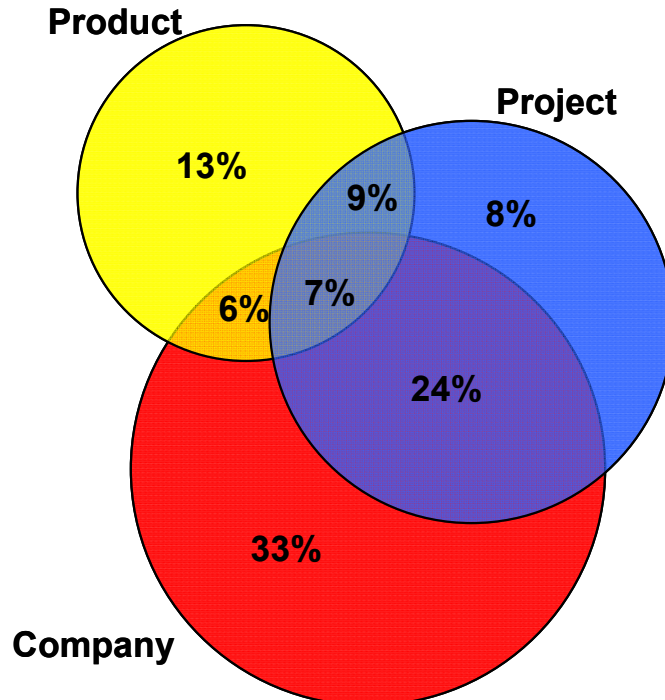
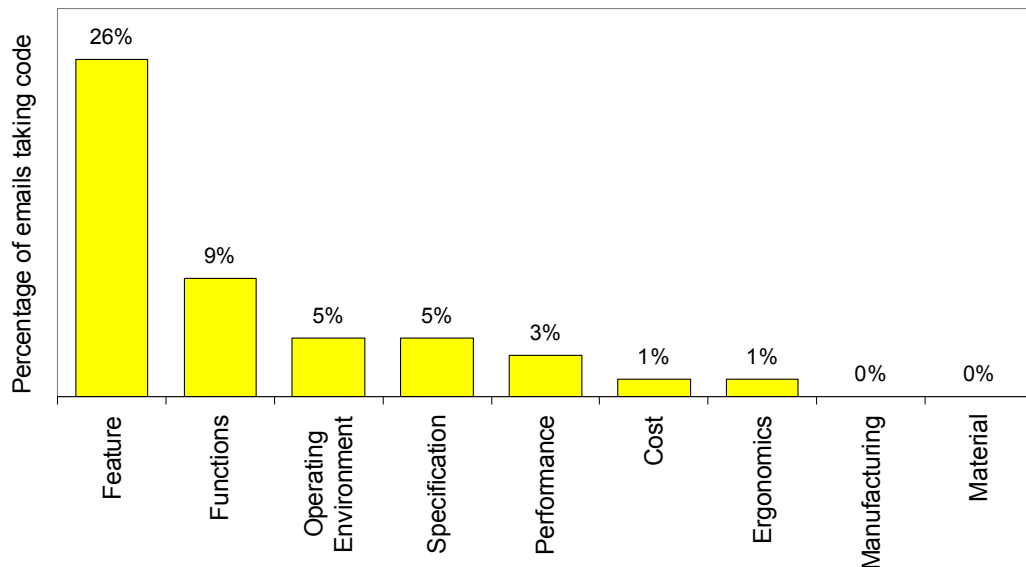


Figure 6-2: The proportions of emails relating to the three topic groups and combinations thereof.

The majority of emails relate to the *Company* group, although many of these emails also involve the *Project and Product*. The prominence of the *company* aspect is reflective of the in-house nature of the project, and its intended application across the company.

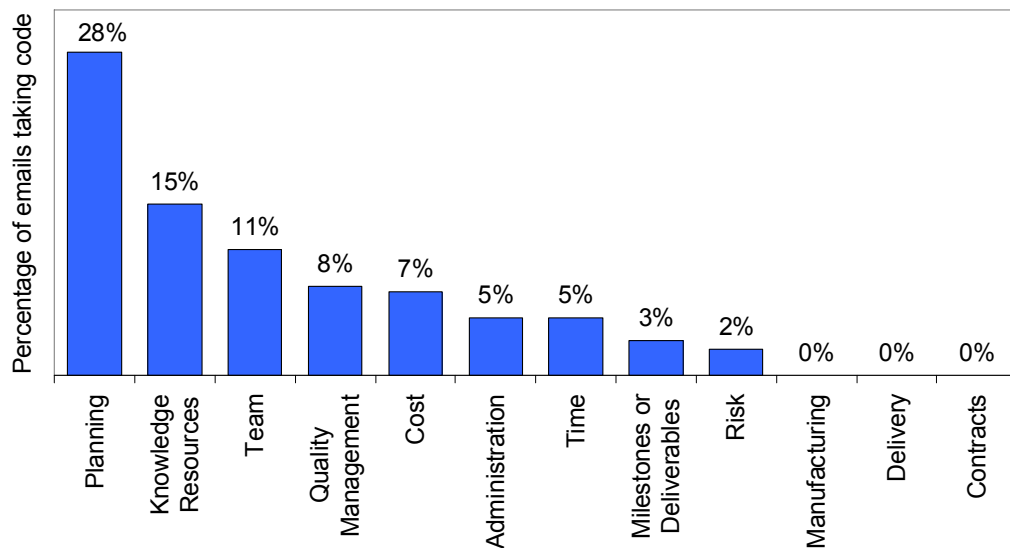


**Figure 6-3: The percentage of emails across the sample coded by terms from within the Product group.**

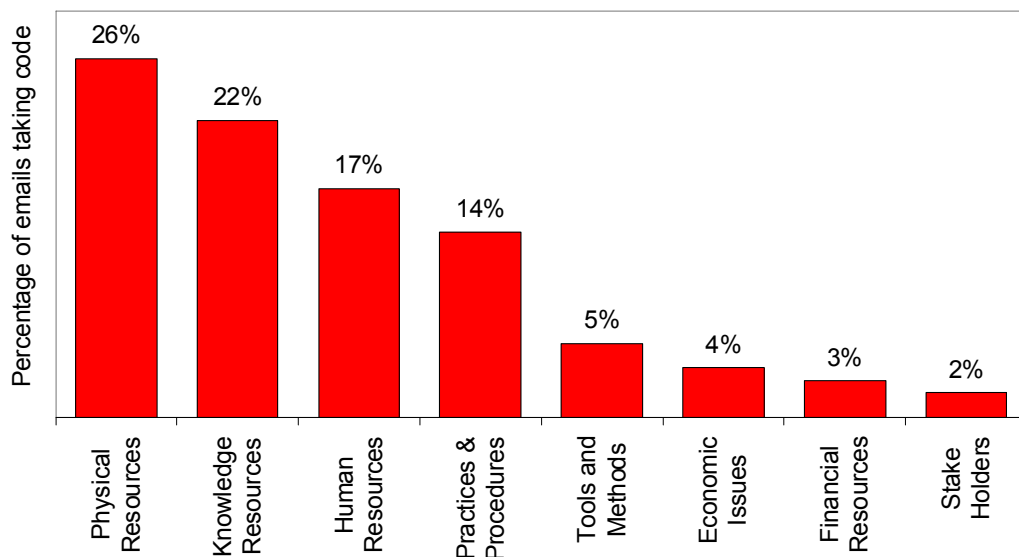
A breakdown of the use of terms which comprise the *Product* group is presented in Figure 6-3. Representing 26% of all the emails sent in the *project*, *features* - the way in which some element of the product design is achieved, were widely discussed. *Functions*, *specifications* and the *operating environment* were also frequent topics of emails. The terms *Manufacturing* and *Materials* were both unused, which corresponds with the nature of the project in software development.

Within the *Project* group, Figure 6-4, the most widely used term was *planning*, which occurred in 28% of correspondence. The Knowledge *resources* of the project were often discussed along with issues relating to the project *team*.

Emails relating to *manufacturing*, *delivery* and *contracts* were all absent from the corpus, none of which would have been expected for an internal project of this nature.



**Figure 6-4: The percentage of emails across the sample coded by terms from within the Project group.**



**Figure 6-5: The percentage of emails across the sample coded by terms from within the Company group.**

The percentage of emails which related to the terms from within the *Company* group are shown in Figure 6-5. The most commonly used term

within this group, *physical resources*, was no more widely used than the most common terms from within the *Product* and *Project* groups; although *Company* related emails were most dominant in the corpus.

In general the company resources aspects *Physical*, *Knowledge* and *Human Resources*, were widely discussed in emails, along with *Practices* and *Procedures*.

### Purposes Emails Were Sent For

The purposes *why* emails were sent were grouped into the three main types, *Information Transactions*, *Management Transactions* and *Problem Solving Contributions*. The prevalence of one or more of these within the emails is shown in Figure 6-6.

The majority of emails were sent for more than one purpose, with *Information* transactions playing a dominant role. Including the overlapping purposes however 30% of emails were still seen to facilitate *Problem Solving* and 34% had a purpose of *Management*.

#### Information Transactions

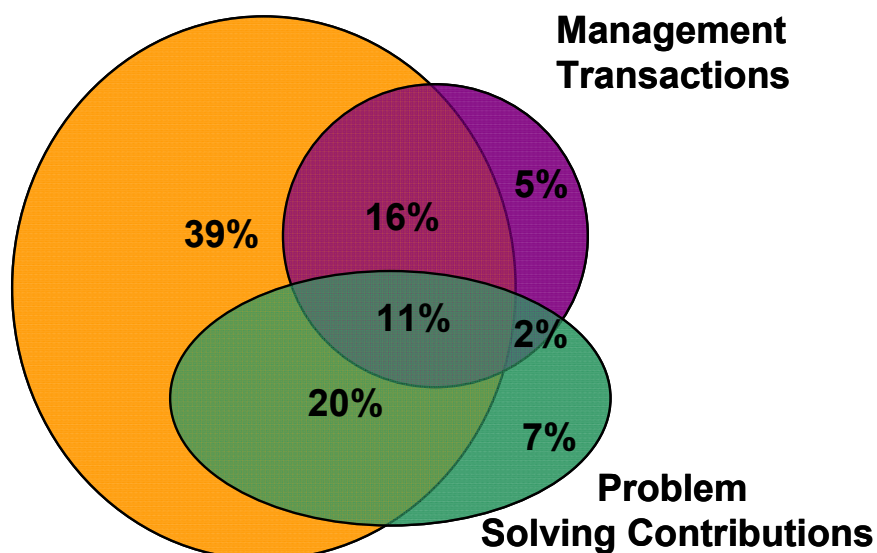
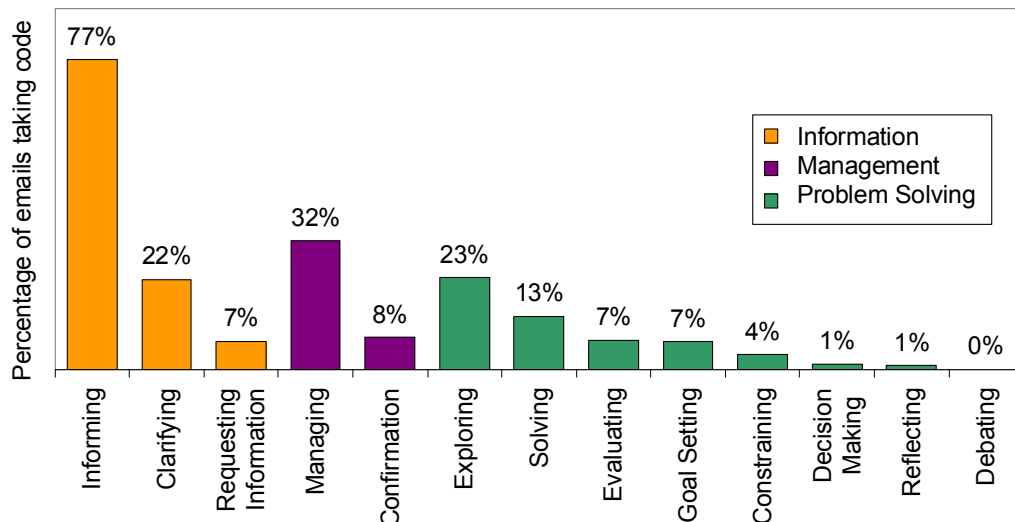


Figure 6-6: The proportions of emails relating to the three purposes for sending emails and combinations thereof.



**Figure 6-7: The percentage of emails across the sample coded by the different elements which comprise the three purposes for sending emails.**

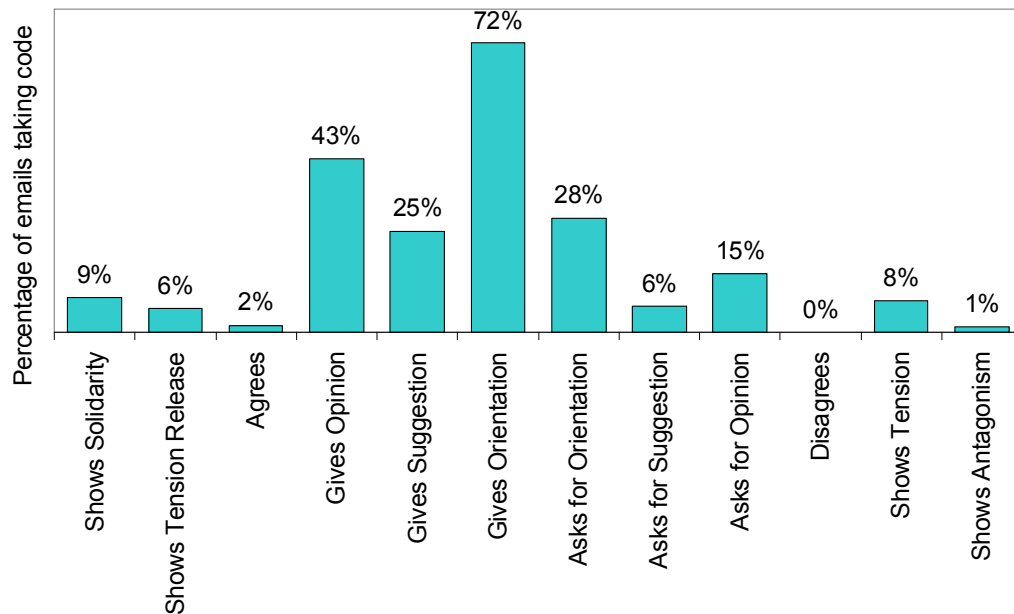
To understand the use of these labels further, a breakdown of the frequencies of the different aspects which comprise the three purposes for sending emails is given in Figure 6-7. It can be seen that when used for *Problem Solving* emails often hosted exploratory discussion (*exploring*), ideas for solutions (*solving*) and evaluations (*evaluating*). *Decision making* and *debating* however rarely took place.

The largest component of *Information* transactions was *informing*, a generic information sharing activity, however *clarification* and requests for information were also often identified in 22% and 7% of emails.

In a similar way *Management* emails were often related to the *Managing* term, but confirmations were specifically identified in 8% of emails.

### How content is expressed

The final part of the coding scheme explored *how* the content of emails was expressed, with each phrase or sentence of an email coded to one of twelve interaction terms. The frequency of each of these is shown in Figure 6-8.



**Figure 6-8: The percentage of emails from across the sample coded by terms which describe how content is expressed.**

The giving of *opinion*, *suggestion* and *orientation* are expressed in a high percentage of emails. These, the task related terms, are also used to a lesser extent by way of asking.

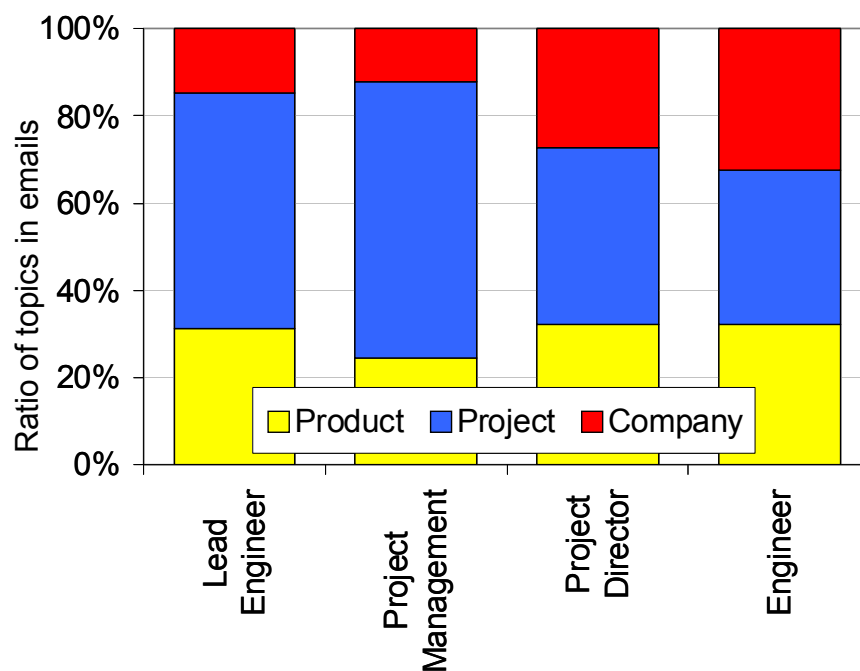
The most widely used emotional terms are *shows solidarity*, in 9% of emails and *shows tension* in 8% of emails. On no occasion was disagreement (*disagrees*) expressed.

#### **6.4.2. People's Use of Email**

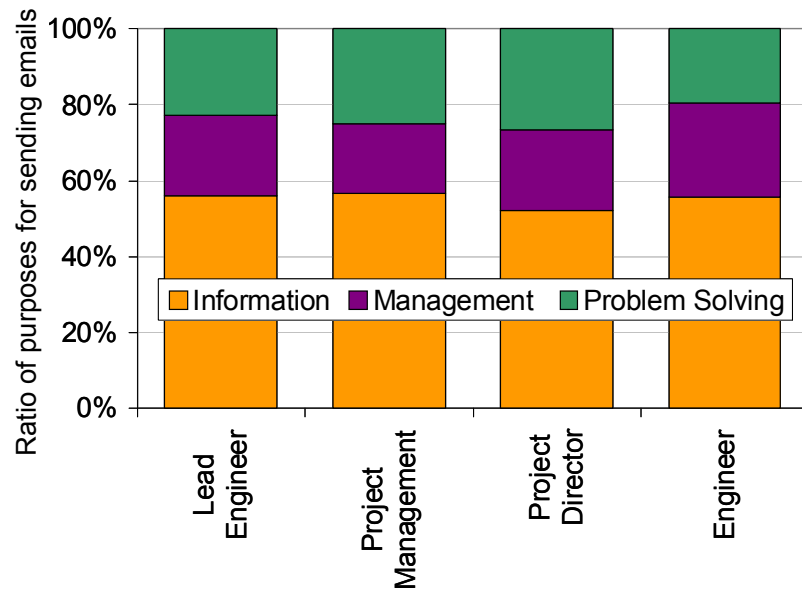
This set of results presents the frequencies with which coding elements were found in emails of different users. This approach enables possible correlations between the different types of user and the nature of their emails use to be identified. In Figure 6-9 and Figure 6-10 the proportions of topics and purposes of emails sent by the four most prominent users in the project are shown.

Figure 6-9 shows that the proportions of topics of discussion do vary considerably between users, although all of the users send emails of each type. The variation in the purposes for which emails are sent, denoted by the three groups, Figure 6-10, is less delineated than the topics which they relate to.

Although the results in this pilot case study show limited insight into correlations between role and email use it is recognised that this may be due to the nature of the project. Most of the roles shown involved similar work, and the four users interacted heavily with each other, hence discussing the same topics. In a more diverse project with outside parties, a more complex social and organisational structure may create more variation in email use between roles.



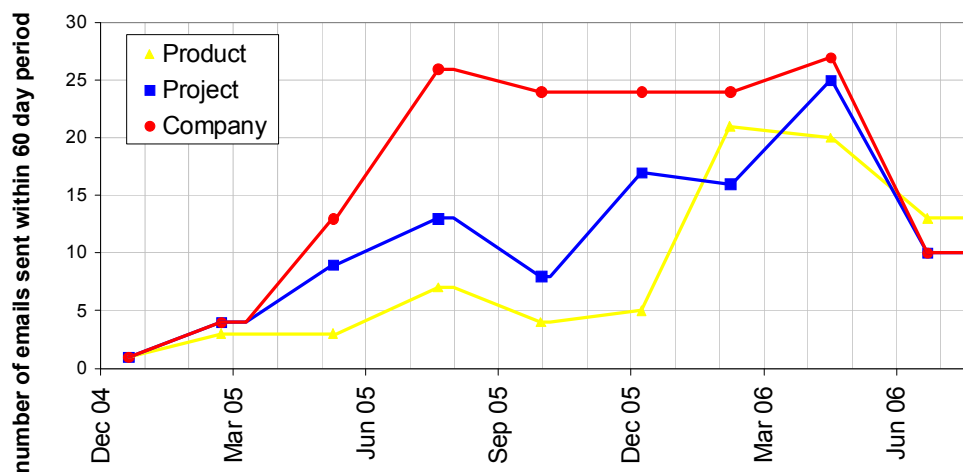
**Figure 6-9: The proportion of Product, Project and Company coded emails sent by four key users in the project.**



**Figure 6-10: The proportion of Information, Management and Problem Solving Purposes coded in emails sent by four key users in the project.**

#### **6.4.3. Changes over the project duration**

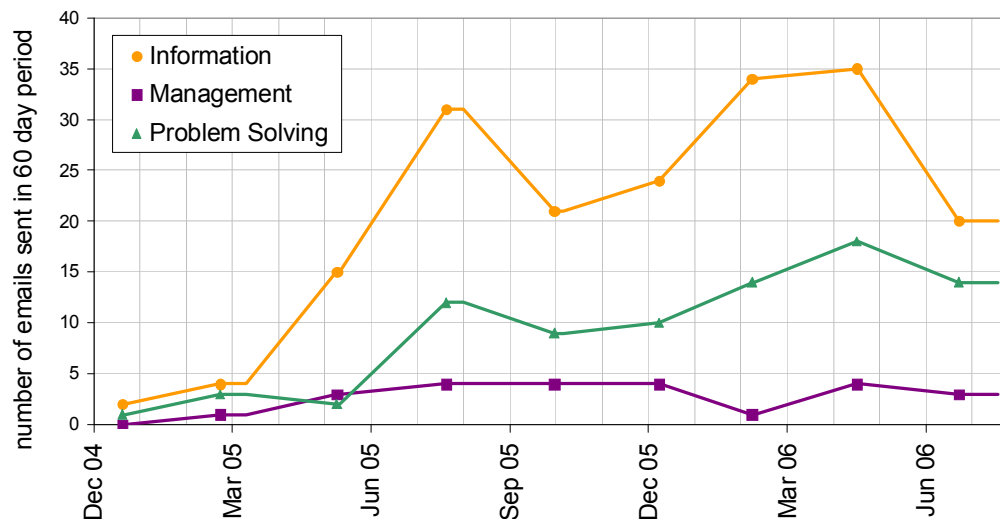
The final set of results presented explores how email use changed over the project duration. In Figure 6-11 the frequency of emails discussing the *Product*, *Project* and *Company* groups during each 60 day period is shown.



**Figure 6-11: The change in the frequency of emails relating to the three topic groups over the project duration.**



It can be seen that emails relating to the *Product* are around three to four times more frequent during the second year of the project than the first. The frequency of *Project* related emails generally rises throughout the project. Emails which relate to the *Company* group grow more rapidly than the other categories, and remain frequent for the continuation of the project.



**Figure 6-12: The change in the frequency of emails relating to the three email sending purposes over the project duration**

Figure 6-12 shows how the purposes for sending emails changed throughout the project duration. The level of *Management* emails remains at a relatively low and stable level throughout the project. The *Information* and *Problem Solving* purposes however are both seen to rise and fall more substantially. For the most these two groups follow a similar trend, with peaks in August 2005 and April 2006. At one point however, May 2005, the two groups trend apart. As an overall observation of the three groups it is noted that the purposes for sending emails do change throughout the project, which may have bearing as an indicator of project progress.

#### **6.4.4. Results Summary**

The results presented have shown the three main analyses possible with the coding approach: frequency of codes and combinations thereof, coding in relation to people's use of email, and coding changes over the duration of the project. A summary of findings from these is now presented.

Most emails in this case study were found to discuss the *company*, followed by the *project* and then the *product*. Each of these types of email occurred in more than 30% of the sample, with many emails relating to more than one of the codes.

Resources of both the *company* and the *project* were extensively discussed in the emails. *Features* were by far the most discussed aspect of the *Product*. A major purpose of emails was found to be the sharing of *information*. Emails were also used to support *Management* and make *Problem Solving Contributions*.

A sample of email users in this case study showed some variation in the purposes and topics which their email sending related to. There was found to be more difference in the topics which emails discussed than the purposes for sending them.

The changing use of email over the project duration was also explored. The frequency of email sending grew over the first few months of the project. Changes were observed over the project duration in the purposes for sending emails and the topics they related to. Based on contextual knowledge of the project (i.e. secondary data) these were suggested to be indicative of project progress.

## 6.5. Chapter Summary

This chapter began to address **Objective 4** of the research: *To apply the proposed approach to recently completed industrial projects and to explore the role that email plays over the project life.*

This was achieved by undertaking and presenting a pilot case study of email use. The primary purpose was to further validate the approach and coding scheme. The secondary aim was to exemplify the potential results which the coding scheme could generate and to draw general conclusions from these which could be considered in the further analysis of the major case study.

The context of the project, an in-house software development project for an engineering company, was outlined. Following which specific detail on how the approach was applied were reported.

The process of coding, data gathering and value of the results attained were then critically discussed. The approach was generally found to be successful, with minor adjustments in the application process recommended for the major case study.

The basic results from the case study were then presented and briefly discussed.

## **7.0 Major Case Study: A large engineering project**

This case study provided an in depth analysis of a real engineering project. The case study is based on a large, complex, multinational project, which is reflective of some of the types of project where email use and long term retrieval and storage are most critical.

Critically the industrial partner for this project was able to cooperate not only in the provision of the email corpus, but also in enabling time with engineers for interviews and discussion and full access to project documents for the secondary analysis.

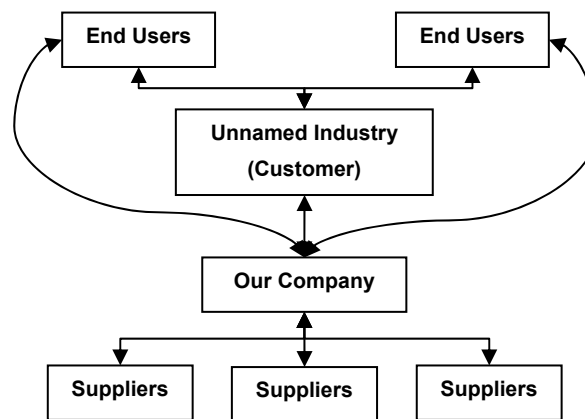
The background for the project in this case study is now discussed with particular detail given to the organisational structure, roles and communication paths. The specifics of how the approach was applied are then discussed. The comprehensive set of results is presented and discussed in relation to the specific project; some conclusions are then drawn.

(Further analysis of selected results in relation to email use in general engineering design settings, beyond the specific case study, is reported in Chapter 8).

### **7.1. Project Context**

The project pertained to a series of six multimillion pound contracts with a foreign unnamed industrial company for the provision of control and power systems to fit six end units. Although these sets (sub-systems and final products) were intended to be replicas of one another, some variations in the designs were required; this was partly due to specific requirements of a number of end users. Examples of this product can be found in marine, defence and aerospace sectors.

The key parties in the project are shown in Figure 7-1. The company was ultimately selling the product to the Heavy Industrial customer with whom their relationship was both highly technical and contractual. However the needs of the end users and future servicing requirements necessitated additional channels of discourse. At the opposite end of the spectrum the company had to deal with numerous suppliers.



**Figure 7-1: Parties involved in the project**

The project was very important to the company, considered to be highly novel and of healthy risk; it was desirable to undertake such a venture to gain understanding prior to other competitors in the field. The company were familiar with the domain of operation and much of the design work involved adaptations and variations to existing sub systems. There were however some technical aspects due to the originality of the design which the company had not dealt with previously and required much consideration. The first unit required the most investment and time, taking three–four years from conception to service with the five following variants lagging by between six months and two years.

Figure 7-2, now discussed, details the relationships between the various roles within the company. The nature of the project required interdisciplinary design from a variety of domains including software, systems, and mechanical engineering. This was undertaken by teams based at two geographically dispersed sites. One site addressed mainly systems and software design and the other mechanical design; each site

had its own Project Manager who was accountable to the Project Director. Additionally a Project Integrator was employed specifically to help align the work of the two sites. A number of sub managers and engineers were answerable to the two Project Managers along with a number of support roles such as administrators and CAD technicians. Commercial and Quality Managers also worked alongside the upper Project Management team. Towards the time of manufacturing, assembly and integration with the customer's product, a site manager was employed and based at this third location.

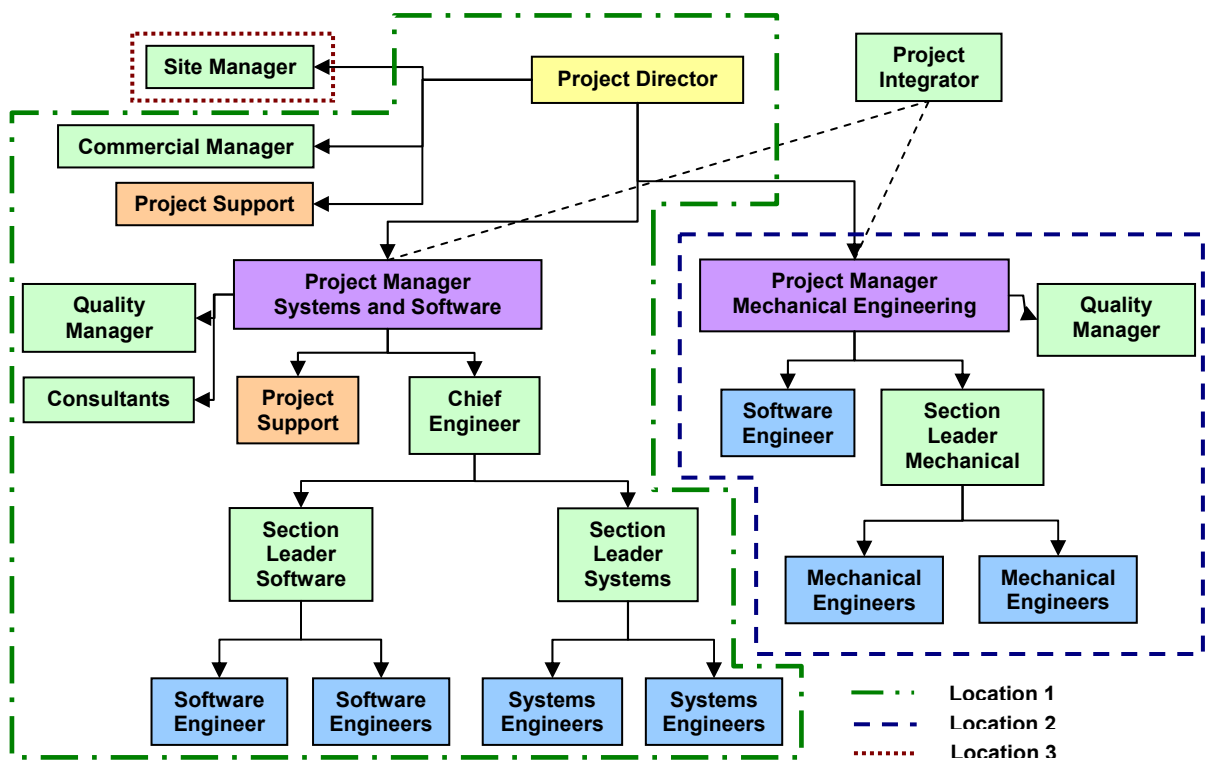


Figure 7-2: The organisation of the project and roles within the company.

The final element contributing to the complexity of the project was the through life service support requirement. This placed special emphasis on reliability and quality in the design and build of the systems, as the company would be required to address issues arising for many years to come. The estimated service life of the products was expected to be over 30 years.

### ***7.1.1. Policy and Practice for Emails, Communication and Documents***

Emails were normally sent to and from individual users personal mail boxes. It was then at their discretion to copy 'relevant' mail into the central repository. It was however possible for users to send mail from the central repository thus creating a stored version in there. The central repository was able to organise emails by features such as date or sender, but also allowed mail to be tagged and sorted into folders. The responsibility for managing and organising the email repository lay with the Project Director.

Shared and managed file systems were also used in the project. These included drawing and software code databases, which supported modifications with version control.

As previously stated the design work was carried out at two geographically dispersed sites. Within each of these, team members were located in a number of offices around the buildings. Hence different project members were able to communicate with one another via *face-to-face* and through internal and external phone calls with varying degrees of ease.

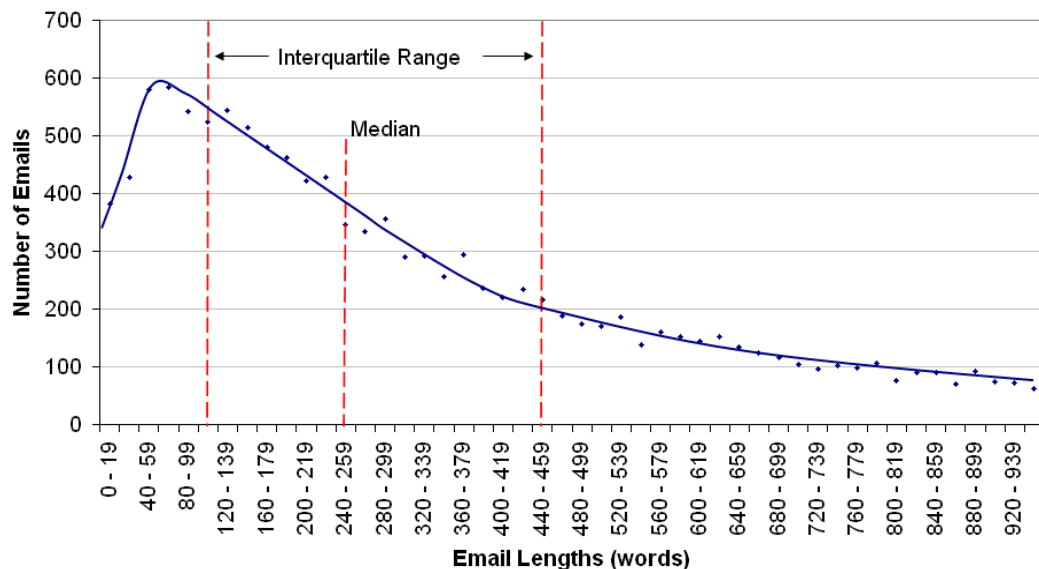
### ***7.1.2. Project Email Corpus***

The email corpus resulting from the project contained around 16 000 emails. These were sent over the first four years of the project, by which point the first product had fully entered service.

The corpus contained emails from a total of 650 senders, around 30 of whom were core team members from the industrial partner. The remainder of contacts included suppliers, end users, the customer and also other workers from within the partner company. From within this base there were a total of 1 080 recipients of email. The number of unique relationships, defined as both the sending and receiving of at least

one email between two users, was 2 150. Thirty two percent of emails were sent between the industrial partner and external parties. An analysis of network and relationships is reported elsewhere by Loftus et., al (2009).

The distribution of email lengths of the corpus is presented in Figure 7-3. The median length of an email was 240 – 359 words. A quarter of emails were less than 120 words, and half of emails were between 120 and 459 words long. Overall three quarters of emails were between 0 and 459 words long. The characteristic of the distribution is similar to the Pilot case study, although the emails tend to be longer by 100 words.



**Figure 7-3: Distribution of email lengths in the Major Case Study corpus**

## 7.2. Application

The entire process of coding and other data gathering was conducted on the premises of the industrial partner, for reasons of data sensitivity. This did however provide the benefit of allowing the research team to gain an impression of the company's modus operandi.



It was recognised in the evaluation of the pilot case study (6.3.3) that fewer emails could have been sampled to achieve a similar result. In order to evaluate a representative sample from the Major Case study, 800 emails were coded between two researchers. These were evenly distributed across the corpus, approximating to the coding of every twentieth email. Of these emails, 60 were coded by both researchers in order to establish the level of inter-coder reliability. Each day the researchers also coded a number of emails together, not included in the sample, to improve their consistency and agreement.

The researchers coded using the *What* and *Why* coding terminologies, previously presented in Table 5-5 and Table 5-6. It was recognised previously that the *How* element of the coding scheme, Table 5-7, increased significantly the length of time taken to code each email. To meet time constraints and enable as large a sample as possible to be coded, the *How* element of the coding scheme was therefore only applied to a smaller sub-set of around 60 emails. As for the pilot study, the software package NVivo (QSR, 2008) was used for the coding of emails.

The coding of each email took an average of 5 minutes, significantly less than the 10 minutes taken in the pilot study. This equated to around six full working days of coding although in practice this was not achieved. The process of coding was intensive and tedious and it was important for coders to maintain a consistent and high quality of work. Coding was therefore interspersed with regular breaks and other activities such as interviewing engineers. There was also a significant amount of time spent addressing the IT logistics of accessing, extracting, formatting and preparing data for coding. For these reasons the total data collection period therefore took a total of four weeks.

A number of formal interviews were conducted with users who were chosen because of their diverse roles in the project, these are summarised in Table 7-1. The interviews generally lasted around one hour, following the semi-structured format reported in Appendix A.

Informal discussions also took place with other project members and company workers, which added to the researchers wider understanding email use and the company.

**Table 7-1: Roles descriptions for the project team members interviewed.**

<b>Persons role</b>	<b>Description</b>
Project Director	Responsible for the entire project
Project Support	Supporting the project director, and responsible for the day to day running of the project. In regular communication with external parties.
Project Manager	Responsible for the systems and software development aspect of the project.
Engineer	Supporting the project manager also interviewed and working in software development.
Commercial Manager	Responsible for resolving contractual disputes, frequently reviewing correspondence histories.
Warranty Support	Responsible for communication with end users over arising issues and arranging engineering support.

Relevant project documentation was reviewed by the two researchers, throughout the data collection period. Documents relating to the structure and roles of the project team were of particular interest. Monthly reports and time plans indicating project activities were also vital. Where possible relevant documents were desensitised and retained for later analysis.

### **7.3. Validation of Coding**

The coding scheme itself, and the overall approach taken had been extensively tested and evaluated as reported in the previous chapters and in particular through the pilot case study.

The inter-coder reliability, statistical significance and stability of the coding scheme were all possible to test quantitatively and are now reported. These confirmed the reliability and validity of the approach.

### **7.3.1. Inter-coder Reliability**

The inter-coder reliability of the two researchers was assessed over a sample of 60 emails which each had coded. Cohen's Kappa (Cohen, 1960) was used to calculate agreement, taking into account chance agreement. For the top level groups coefficients above the suggested benchmark (0.7) for an exploratory study were attained (Lincoln and Guba, 1985), as shown in Table 7-2. For the vast majority of terms within these groups coefficients  $>0.7$  were found, with the exceptions being lesser used terms.

**Table 7-2: Inter-coder reliability of top level codes  
calculated using Cohen's kappa.**

<b>Code</b>	<b>Coefficient</b>
Product	0.87
Project	0.85
Company	0.81
Information Transactions	0.82
Management Transactions	0.76
Problem Solving Contributions	0.71

Not only were suitable kappa coefficients achieved, but by coding alternate emails across the sample, inter-coder inconsistencies were less likely to bias results examining the change in frequency of codes over the project duration.

### **7.3.2. Statistical Significance**

The statistical significance of each code was established using the split half technique, which compares the difference in frequency of codes in each half of the sample. This is translated to a coefficient where 1 is significant and 0 not significant.

**Table 7-3: Statistical Significance of Top Level Codes Using the Spilt Half Technique**

<b>Code</b>	<b>Coefficient (1- <math>\alpha</math>)</b>
Product	0.95
Project	0.91
Company	0.99
Information Transactions	0.88
Management Transactions	0.82
Problem Solving Contributions	0.79

As is shown in Table 7-3 the significance of all of the top level terms was seen to be high ( $>0.79$ ). The significance of terms within each of these groups was also calculated, with an average value (0.51). This was recognised to be seriously reduced by several terms which are used so infrequently, for example *Company Financial Resources*, that their significance cannot realistically be determined. The median value (0.79) provides a more meaningful and constructive measure.

### **7.3.3. Stability**

To verify the stability of the coding scheme, two batches of 30 emails, each coded five days apart were considered. As shown in Table 7-4 acceptable coefficients ( $>0.78$ ) were shown for all of the top level terms applied. For the terms within these groups median values ( $>0.67$ ) were agreed to be acceptable given the reduction in significance due to the limited sample size.

**Table 7-4: The stability of top level codes checked five days apart**

<b>Code</b>	<b>Coefficient (1- <math>\alpha</math>)</b>
Product,	0.99
Project	0.96
Company	0.95
Information Transactions	0.99
Management Transactions	0.84
Problem Solving Contributions	0.78

## **7.4. Results**

The results of the coding of the email corpus are now presented with an initial discussion. Further analysis, particularly relating to the supplementary data sources is presented in Chapter 8. The first section of results presents the frequencies of different elements and combinations of elements of the coding scheme across the sample. The second section then examines how coding varied between different emails users. The final section then presents findings of how the coding changes over the project duration.

### ***7.4.1. Frequency of Coding Elements***

As the most fundamental results, these present the frequency with which each element of the coding scheme was applied to an email. The percentages given in this section correspond to the frequency of coding elements across the entire sample (unless otherwise stated). This can be viewed as representative of the percentage expected across the entire corpus, given that the results were shown to be statistically significant.

The frequency of topics which emails related to are first presented, followed by the frequencies of the various purposes for which they were sent. The frequency of combinations of topic and purpose are then presented in order to consider possible correlations. Finally the frequencies of codes for how content is expressed are presented.

### Topics Which Emails Contained

Figure 7-4 depicts the proportions of the three topic groups, *Product*, *Project* and *Company*, which emails related to. This also shows the combinations of these categories, where emails contained more than one topic of discussion.

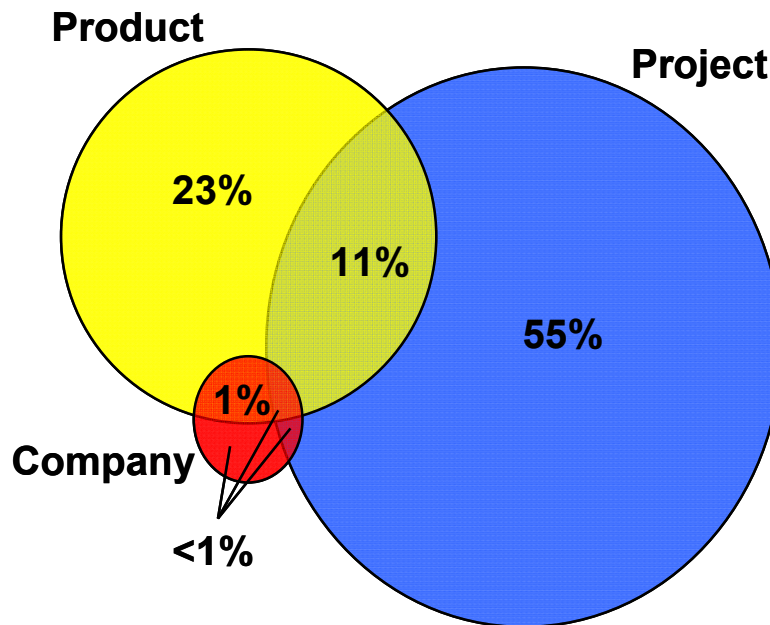
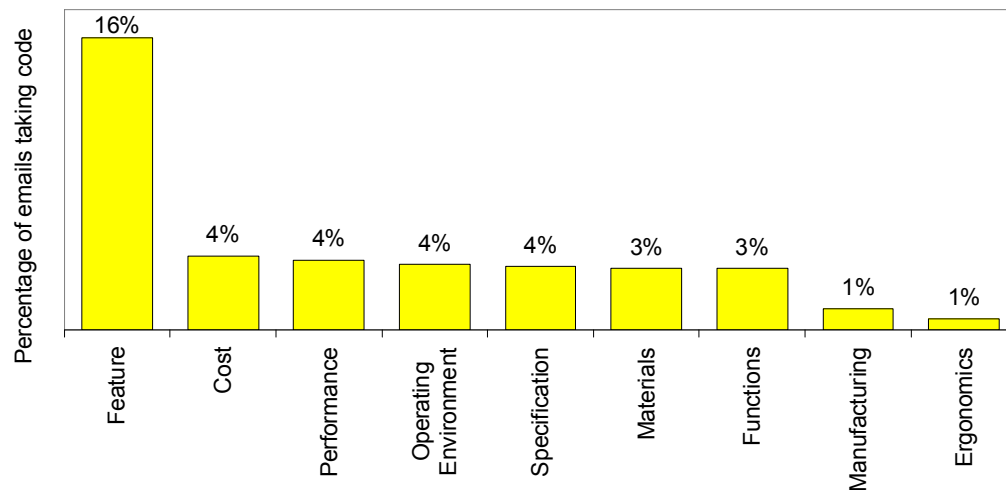


Figure 7-4: The proportions of emails relating to the three topic groups and combinations thereof.

The majority of email content, over 60%, related to the *Project*. Twenty three percent of emails related solely to the *Product* and its design, with over a third of emails including this as well as another aspect. A very small proportion of emails, less than 4% related to the *Company* in some way.

Figure 7-5, Figure 7-6 and Figure 7-7 show the use of the terms within each of the three main groupings (*Product*, *Project* and *Company*). The percentages relate to the use of the term within the entire sample, rather than within the respective grouping. As an emphasis to the significance of these results, it is noted that a term which occurs in just one percent of

the sample represents in the order of 160 emails, within the project corpus.

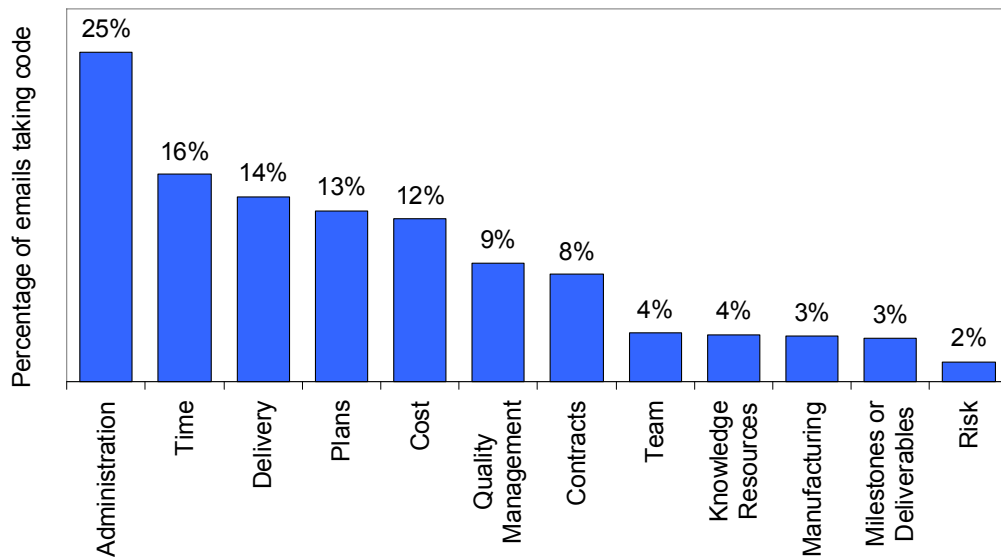


**Figure 7-5: The percentage of emails across the sample coded by the terms within the Product group.**

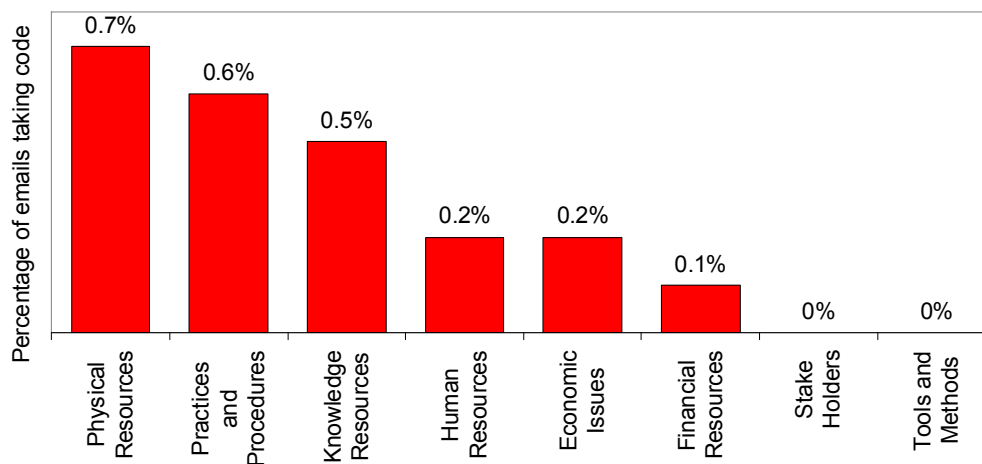
Figure 7-5 shows the terms used from within the *Product* group. Emails were most often used to discuss *features*, 16% of emails. This is evidence for a key part of design, the generation of different possible ideas for achieving something. Aspects which these features might be considered in relation to, *cost*, *performance*, *materials*, *specification* each occur in 3 or 4% of emails. The small proportion of *manufacturing* information, 1%, is in line with expectations, given that this would have been more of a concern to suppliers from whom components were bought. The comparatively infrequent discussion of *ergonomics* is also logical given that the project focuses on systems integration, of which the human interface forms a small component.

Emails were most often coded to the *Project* group of terms, the breakdown of which is given in Figure 7-6. It was found that a quarter of emails are of an *administrative* nature, these include tasks such as arranging meetings, booking facilities and passing on messages. Sixteen percent of emails also contain some reference to *time*. A major role of emails in this project was to facilitate *planning*, 13% of emails. Although only 2% of emails were coded with *risk*, this could equate to around 320 emails during the project which highlight a possible problem which has

yet to occur. Just one of these may contain a piece of information whose effective distribution may make or break the project.



**Figure 7-6: The percentage of emails across the sample coded by terms from within the Project Group.**



**Figure 7-7: The percentage of emails across the sample coded by the terms from within the Company group.**

A small proportion of the emails in the corpus related to the *Company* category; the breakdown for which is shown in Figure 7-7. The most common aspect of these was *physical resources*, occurring in just 0.7% of the sample this still suggests that over 100 emails in the project corpus were of this nature. Company *practices and procedures*, *knowledge*,

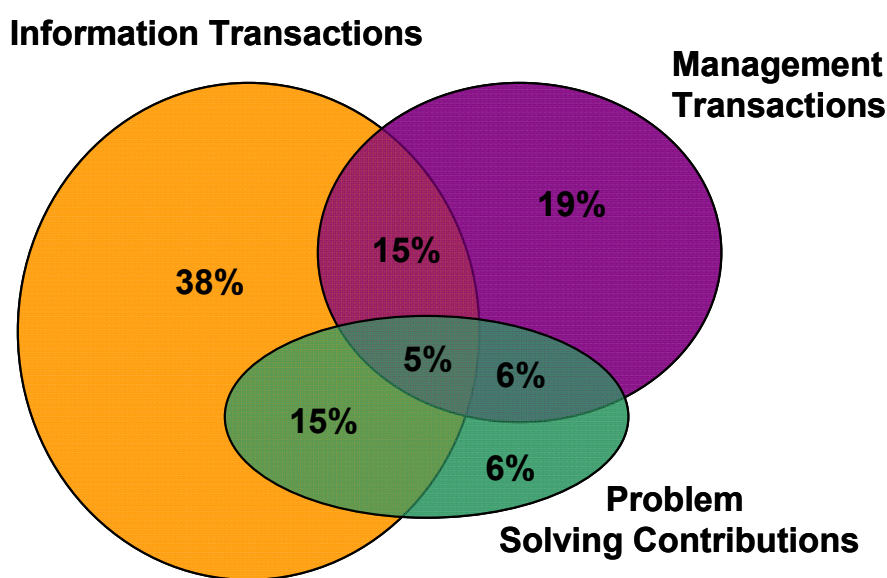


*human and financial resources* were also discussed along with *economic issues*.

### Purposes Emails Were Sent For

Figure 7-8 depicts the proportions of the three groups, *Information Transactions*, *Management Transactions* and *Problem Solving Contributions*. Each email was found to fulfil one or more of these purposes, with intersections also shown in the figure.

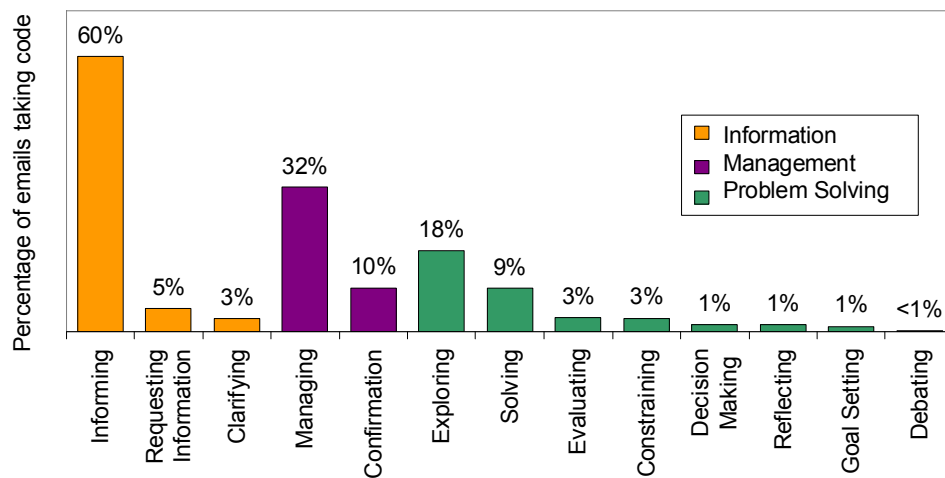
Over 70% of emails were sent with the purpose of an *information transaction*, with 38% of emails sent only for this purpose. Six percent, equivalent to nearly 1 000 emails, were sent purely for the purpose of *problem solving*. Around a quarter of emails had a *problem solving* purpose and also a further *information* or *management* purpose.



**Figure 7-8:** The proportions of emails relating to the three purposes for sending emails and combinations thereof.

Figure 7-9 shows the frequency across the sample of the individual terms, which comprise the three overall purposes for sending emails. The most dominant of these is *Informing*, occurring in 60% of emails. *Requesting information* and *clarifications* occur far less often, but still in around 800 and 500 emails respectively. A third of the emails manage

and direct activity (*managing*), and notably one in ten emails requests or provides confirmation (*confirming*).



**Figure 7-9: The percentage of emails across the sample coded by the different elements which comprise the three purposes for sending emails.**

*Exploring* and *Solving* were the two most common forms of *problem solving* discussion. The aspects of *constraining* and *evaluating*, although less frequent, still each occur in around 500 emails in the project.

### Combinations of Coding

As previously shown, the vast majority of emails in the corpus related to the *Product* and *Project* categories, with few discussing the *Company*. Hence from herein only the former two categories are considered in the results reported.

Having determined the frequency of *Product* and *Project* topics of discussion, and the purposes for which emails were sent (*Information*, *Management* and *Problem Solving Contributions*) it was then examined if there was a correlation between these. The percentage of emails for each combination are listed in Table 7-5.

**Table 7-5: The percentage of emails for each combination of purpose and topic.**

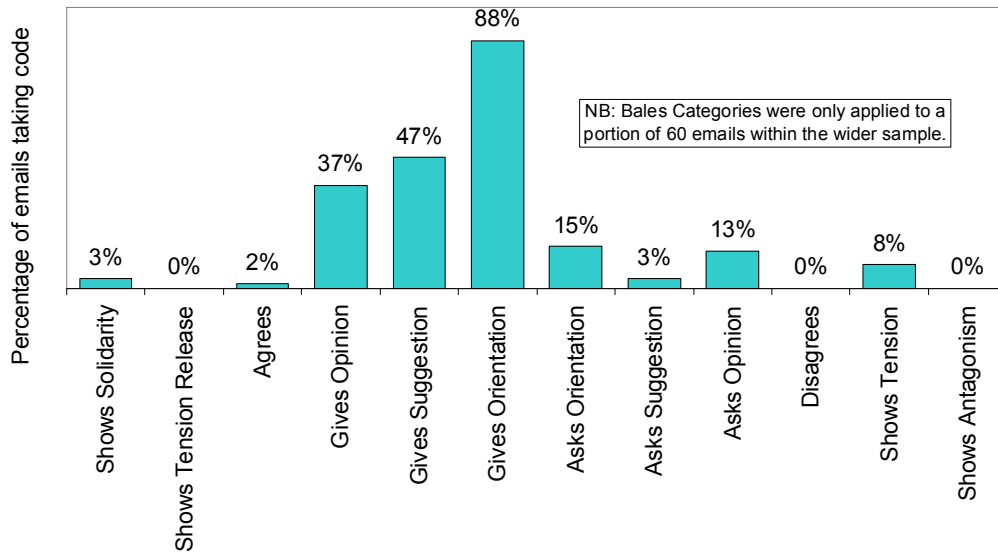
	Only Information	Only Management	Only Problem Solving	Only Information and Management	Only Information and Problem Solving	Only Management and Problem Solving	All three Purposes
Only Product	7.6	3.2	2.5	2.1	4.7	2.8	1.9
Product & Project	5.3	3.2	0.3	1.7	1.6	0.1	0.8
Only Project	25.0	12.8	3.4	11.0	5.4	2.6	2.1

The most frequent combination, 25% of emails, had the sole purpose of *information* transactions in relation to the project. A further 12.8% had the sole purpose of facilitating *Management* in relation to the *Project*. The third most common type of email, 7.6% were *information* transactions about only the *Product*. Thirty four percent of emails contained aspect of both a *management* purpose and a *project* topic.

### **How content is expressed**

A smaller portion of the sample was coded using the terms to describe how content was expressed, derived from Bales (1950) interaction categories. Each phrase within each email was coded to one of these categories. The bars in Figure 7-10 show the percentage of emails which contained at least one occurrence of each code.

Eighty-eight percent of emails contained *Giving Orientation*, 47% *Giving Suggestion* and 37% *Giving Opinion*. A far smaller proportion of emails asked for *Orientation*, *Suggestion* and *Opinion*, 15%, 3% and 13% respectively. The emotional interactions, showing *tension*, *agreement* and *solidarity* occurred in 8%, 2% and 3% of emails.



**Figure 7-10: The percentage of emails from across the sample coded by terms which describe how content is expressed.**

#### **7.4.2. Variation in Users**

This section of results explores the differing use of email between users. From the 650 project participants to have sent emails, only the major senders, for whom the data is sufficiently representative and meaningful are included in the following results. These 27 users represent a spectrum of roles and parties in the project.

The topics of emails from different users are first discussed, followed by the purposes for which they sent email.

#### **What topics different users discuss**

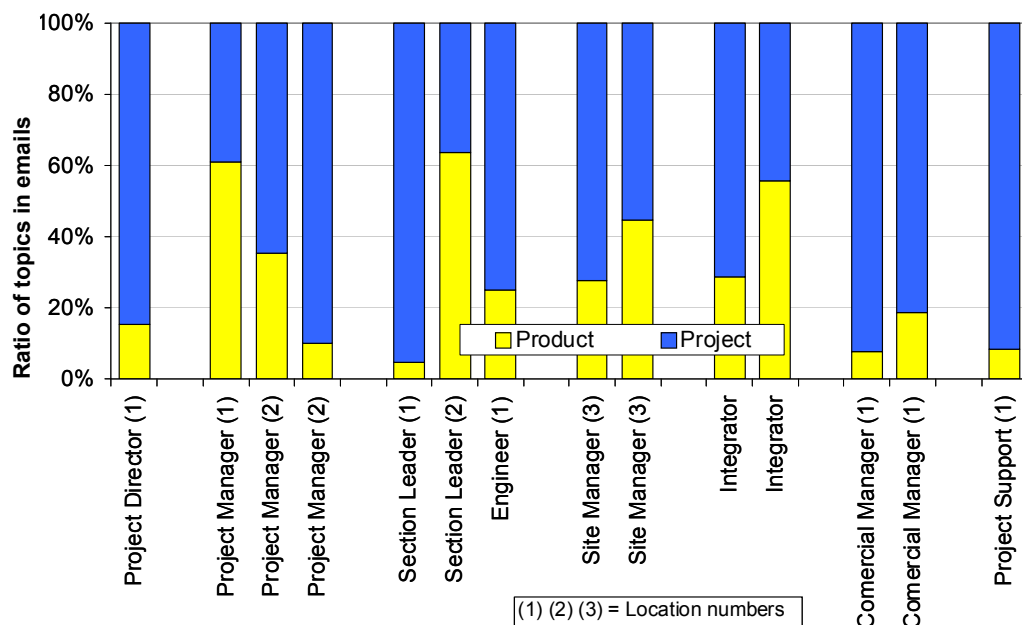
Common wisdom would suggest that users who have different roles would send messages of differing content. To examine this, the ratio of *Product* to *Project* coded emails sent by different users is considered; because of its infrequent use the *Company* group is not included.

Figure 7-11 shows the ratio of *Product* and *Project* related emails for key users within the industrial partners company. All but three of the 14 users

sent more *Project* than *Product* type emails with those who did not, sending around 60% *Product* emails.

Lead Engineers and managers tend to use email to discuss the *product* rather than *project* more so than the Project Director. However there is significant variation within each of these groups.

The commercial manager and project support workers, who are not engineers, sent relatively little email which relates to the *product*.



**Figure 7-11: The proportion of Product and Project coded emails sent by key users in the industrial partner company.**

Figure 7-12 shows this data again, but also includes further users from industrial partner and key users from the suppliers, customer and end user. Engineers and Project Managers from the external companies have stronger tendencies to use email to discuss the *product*, in some cases exceeding 90% of their communication. However a number of external engineers and managers still send heavily *Project* dominated correspondence.

Users with similar roles in the same company often show large variations in the ratio of *Product* to *Project* emails they send.

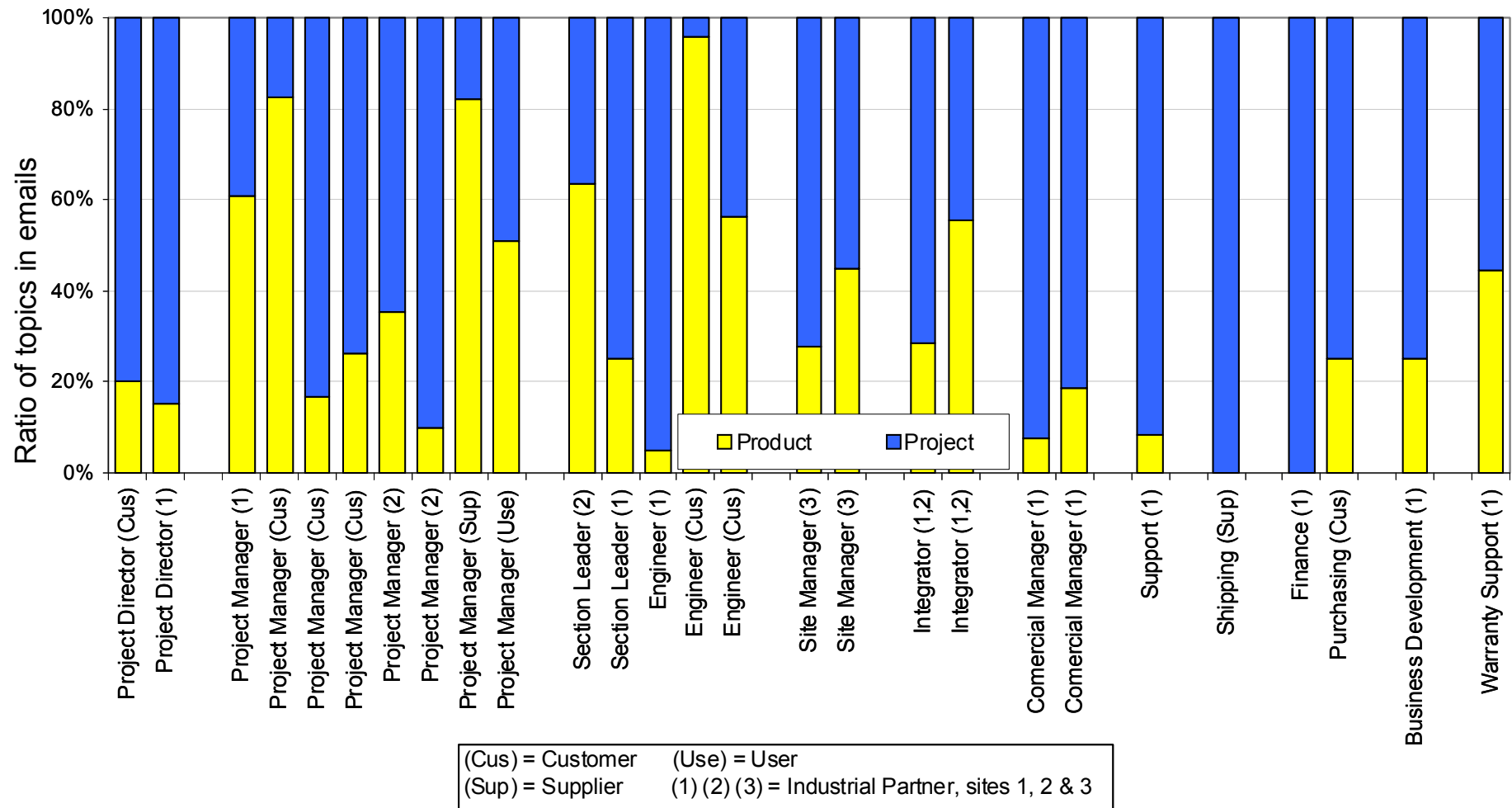


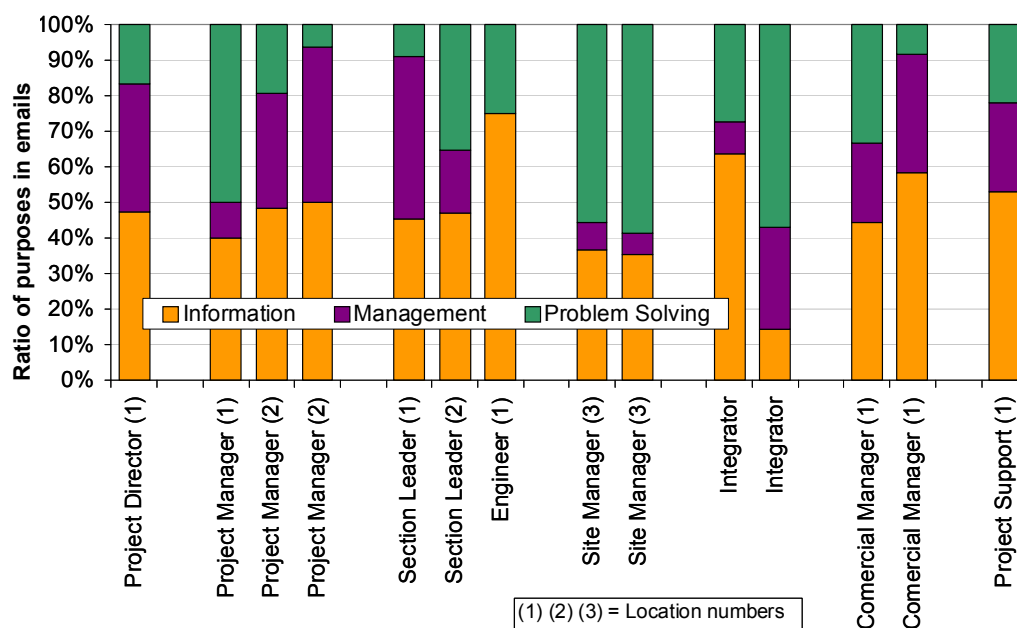
Figure 7-12: The proportion of Product and Project coded emails sent by key users from major parties in the project.

## What purposes for sending email users have

The following results examine whether a users role influenced the purposes for which they sent emails. The proportion of emails coded to the three purposes for sending emails, *Information Transactions*, *Management Transactions* and *Problem Solving Contributions* were considered for the same key groups of users.

Figure 7-13 shows the purposes for users within the industrial partners company. Some limited tendencies for the different roles can be seen.

The Project Director is one of the greater users of email for facilitating management. This is shown similarly for other Lead and Managing Engineers. Notably the non-managing engineer in Figure 7-13 does not use email for the purpose of *management*.



**Figure 7-13: The proportion of Information, Management and Problem Solving purposes coded in emails sent by key users in the industrial partner company.**

The biggest problem solvers all work in engineering roles, the site managers, integrator and a project manager. However one of the

Commercial Managers uses email over 30% of the time as a problem solving tool.

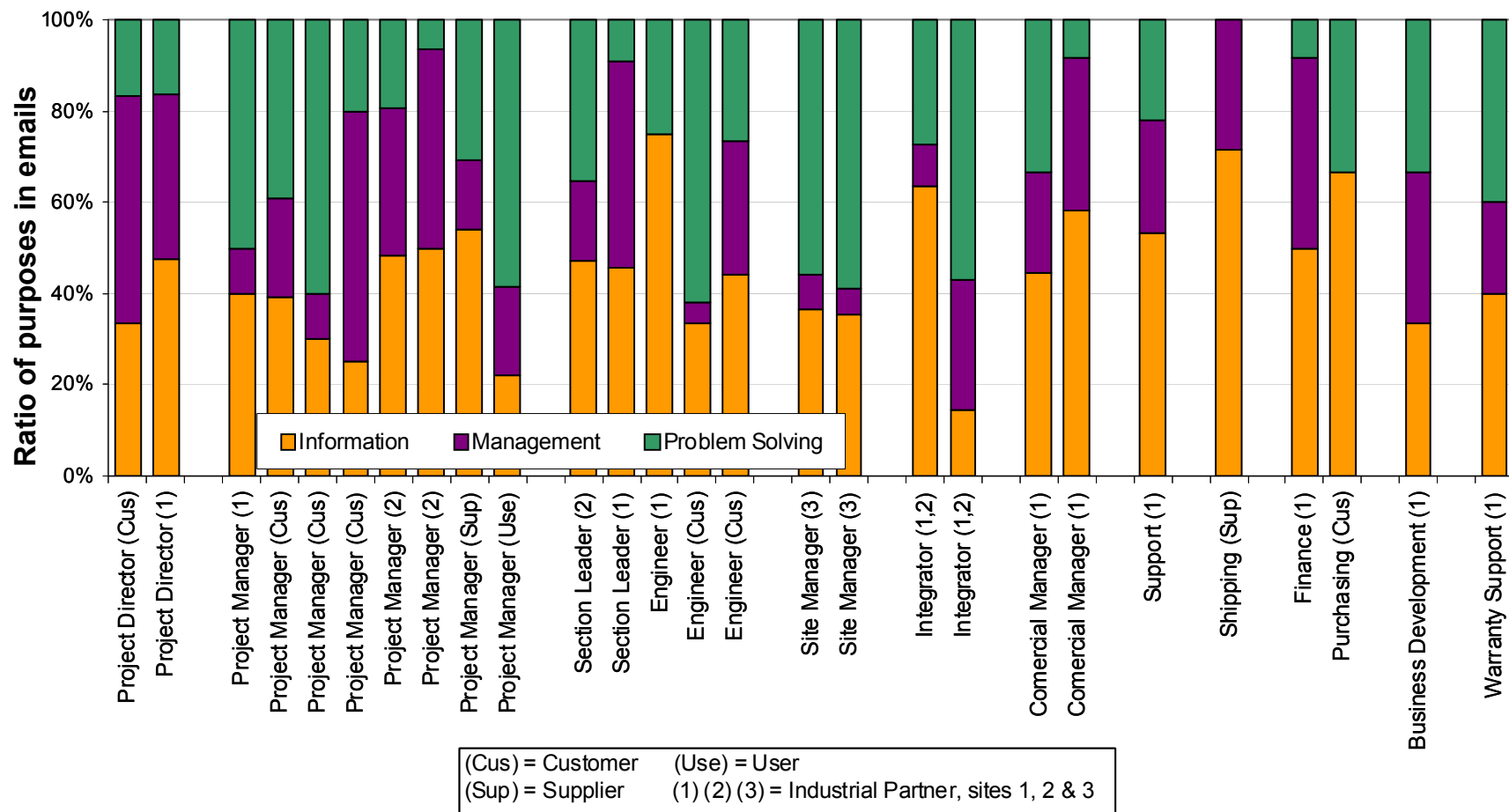
The proportion of emails sent for *Information* transactions is greatest for the Engineer and one of the Integrators, both of whom use email relatively little for management. Otherwise the purpose of information sending does not appear to correspond with any particularly role.

Figure 7-14 shows transactions of key users from the industrial partner, suppliers, customer and end users. Those in engineering roles, including managers, leaders, integrators all show a wide variety of profiles in their email sending purposes. Both project directors however, from the industrial partner and customer are seen have similar profiles.

Those in engineering roles for the end user and customer, show a tendency to use email more for problem solving and management than users from the Industrial Partner company.

Generally it is seen that the purposes for which emails were most often sent varied greatly between users. Factors beyond a user's role and company must explain this.





**Figure 7-14: The proportion of Information Transactions, Management Transactions and Problem Solving Contributions coded in emails sent by key users from major parties in the project.**

### **7.4.3. Change over project duration**

The section of the results explores whether the topics which emails discussed and the purposes for sending them varied throughout the duration of the project.

#### **Changes in the topics which emails contained**

Figure 7-15 shows the number of emails falling into the *Product*, *Project* and *Company* categories during each 60 day interval of the project. The major phases (e.g. manufacture, testing) for the first in the series of six products are shown in the figure. The phases for the further five variant products which were lagged by between six to 18 months are not shown for reasons of clarity, but also because it is known that the majority of email discussion pertained to the first product.

Initially the proportion of emails relating to the *Product* and *Project* is similar. After the first year, when the specification phase has been completed, the topics of emails are more often related to the *Project*.

There is a notable spike in *Product* related emails during the testing phase in March to April 2007, whilst at the same time *Project* related emails remain relatively constant.

The peak for *Project* emails occurs at earlier in the project, suggesting that the pace of this is picking up.

The frequency of *Company* related emails remains low and relatively constant throughout the project, if a little higher during the first two years.

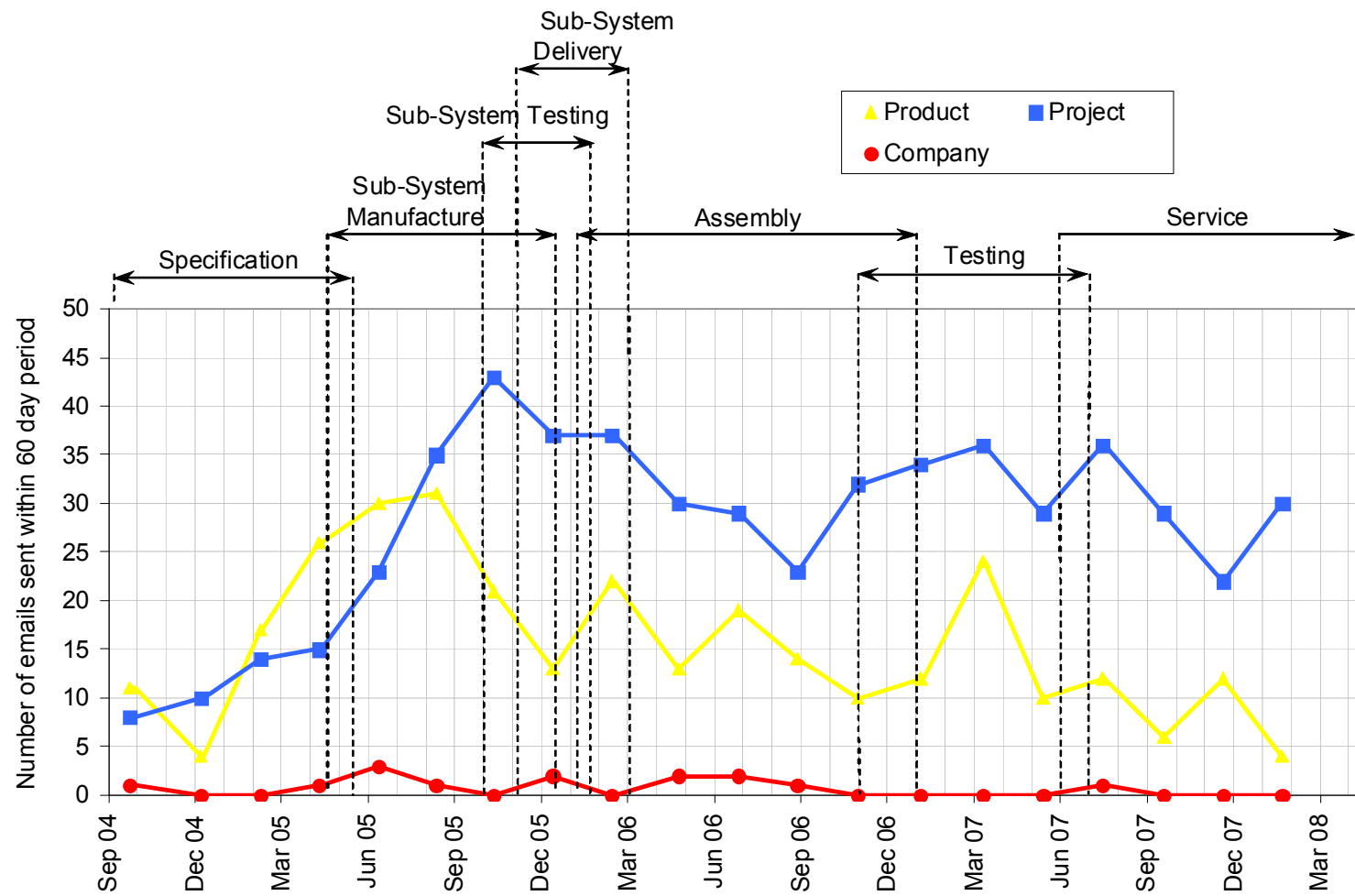
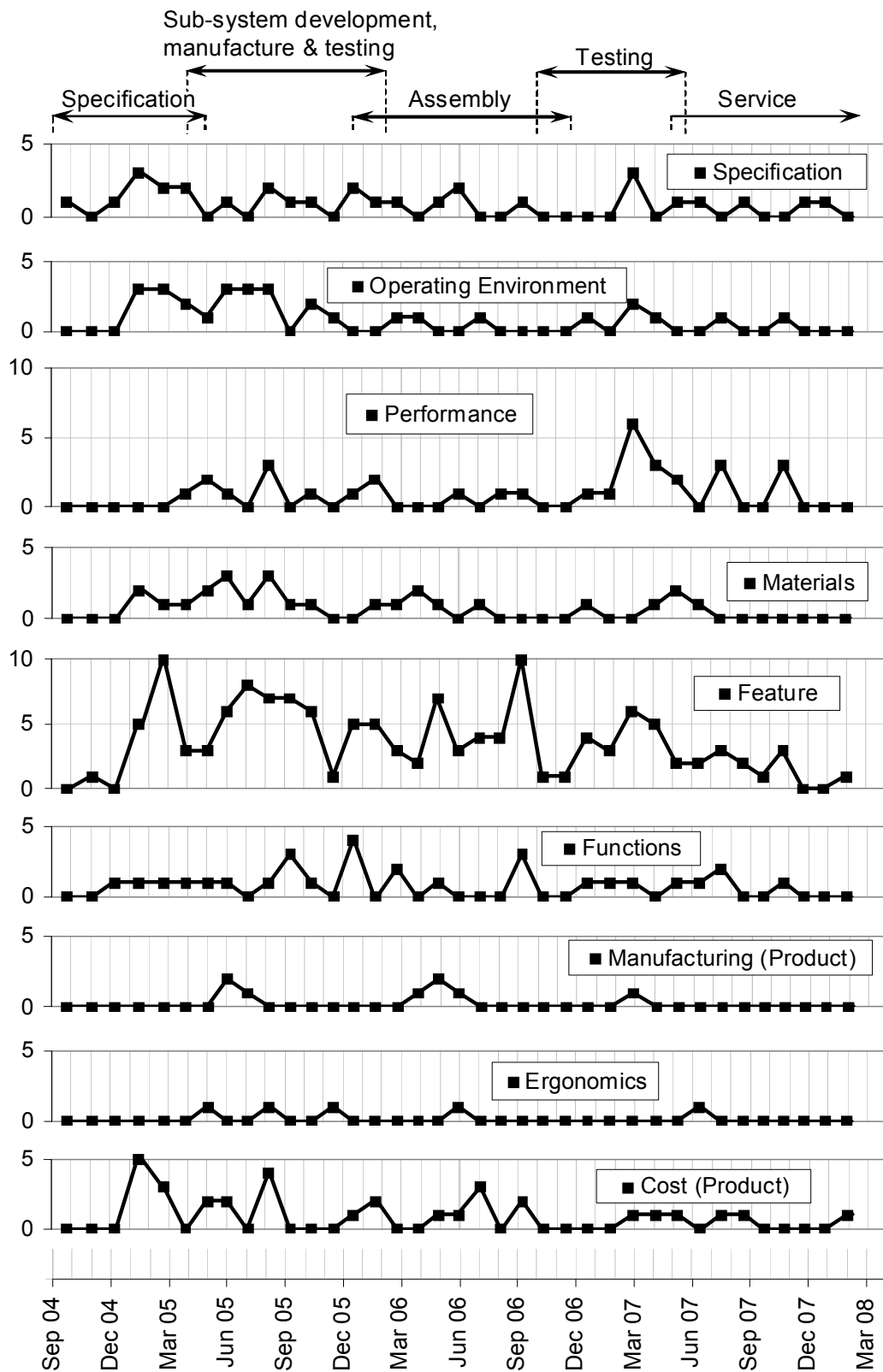


Figure 7-15: The change in the frequency of emails relating to the three topic groups over the project duration.



**Figure 7-16: The change in frequency of emails coded to terms within the Product group over the project duration.**

The use of the various terms within the *Product* group throughout the project duration is shown in Figure 7-16.

Although the use and hence sample size for many of the terms is small the results may still be indicative of trends of use. In particular, *Features* are discussed more in the early stages of the project. *Performance* however is discussed more in the later stages. This may be in the context of verification of performance during testing stages, but might identify performance issues which should have been considered earlier.

The *Functions* are discussed most frequently in the middle stages of the project; this is initially surprising as they should have been specified early on. Both the *Costs* of the product and the *operating environment* occur more in earlier emails, suggesting that they have been considered in good time. The *specification* is a regular topic of discussion in emails throughout the project, which may fit expectations as it should be revisited throughout design work.

A similar breakdown of the use of terms within the *Project* category is given in Figure 7-17 and Figure 7-18.

*Quality Management* is a topic of emails throughout the project but particularly during the later half of 2005. *Contractual issues* show a similar pattern but peak later in February 2006.

Emails which discuss issues of *Time* rise and fall substantially in comparison to other terms throughout the project, with a low point in September 2006. *Project Costs* follow a pattern showing similar variation.

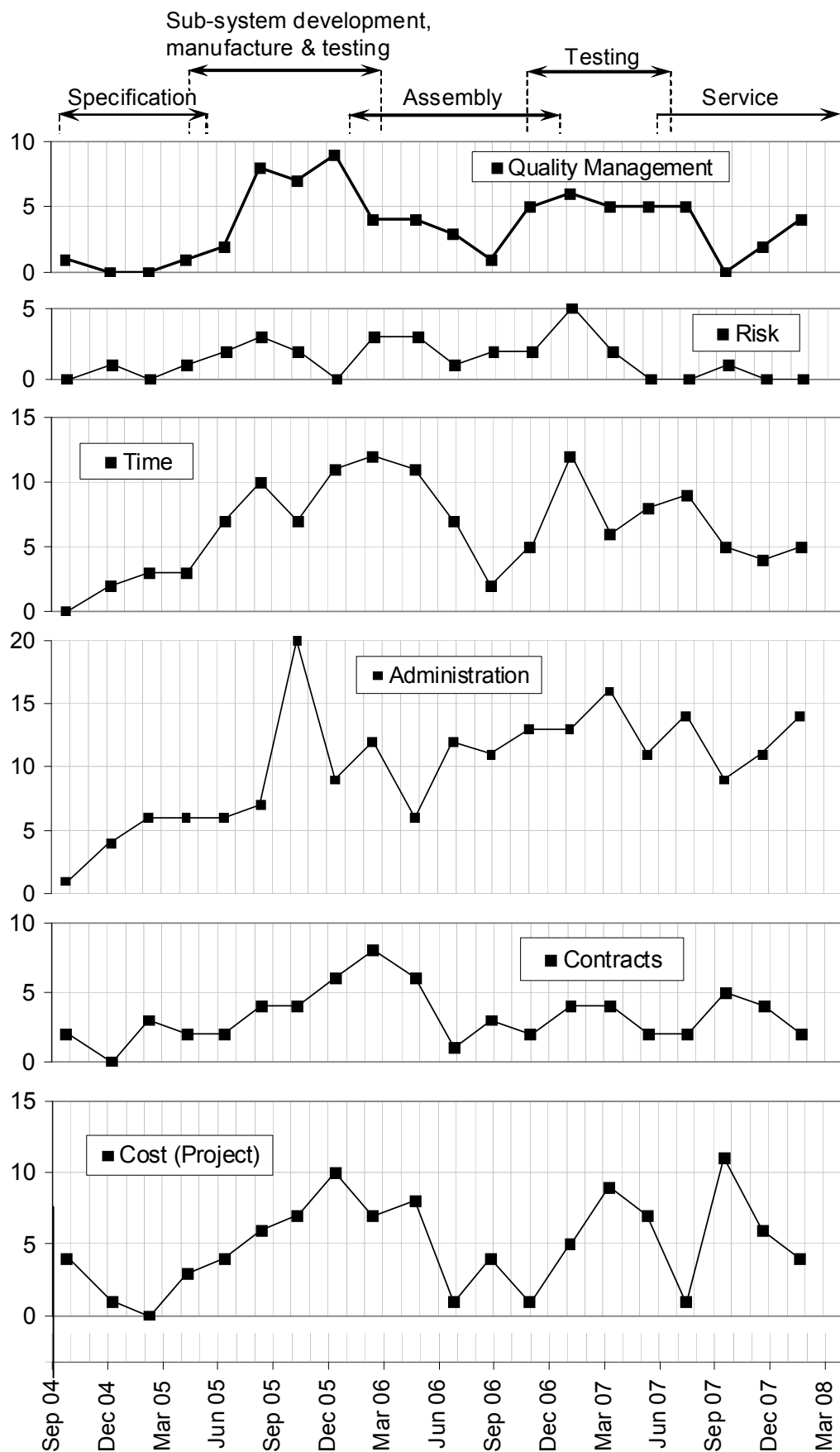
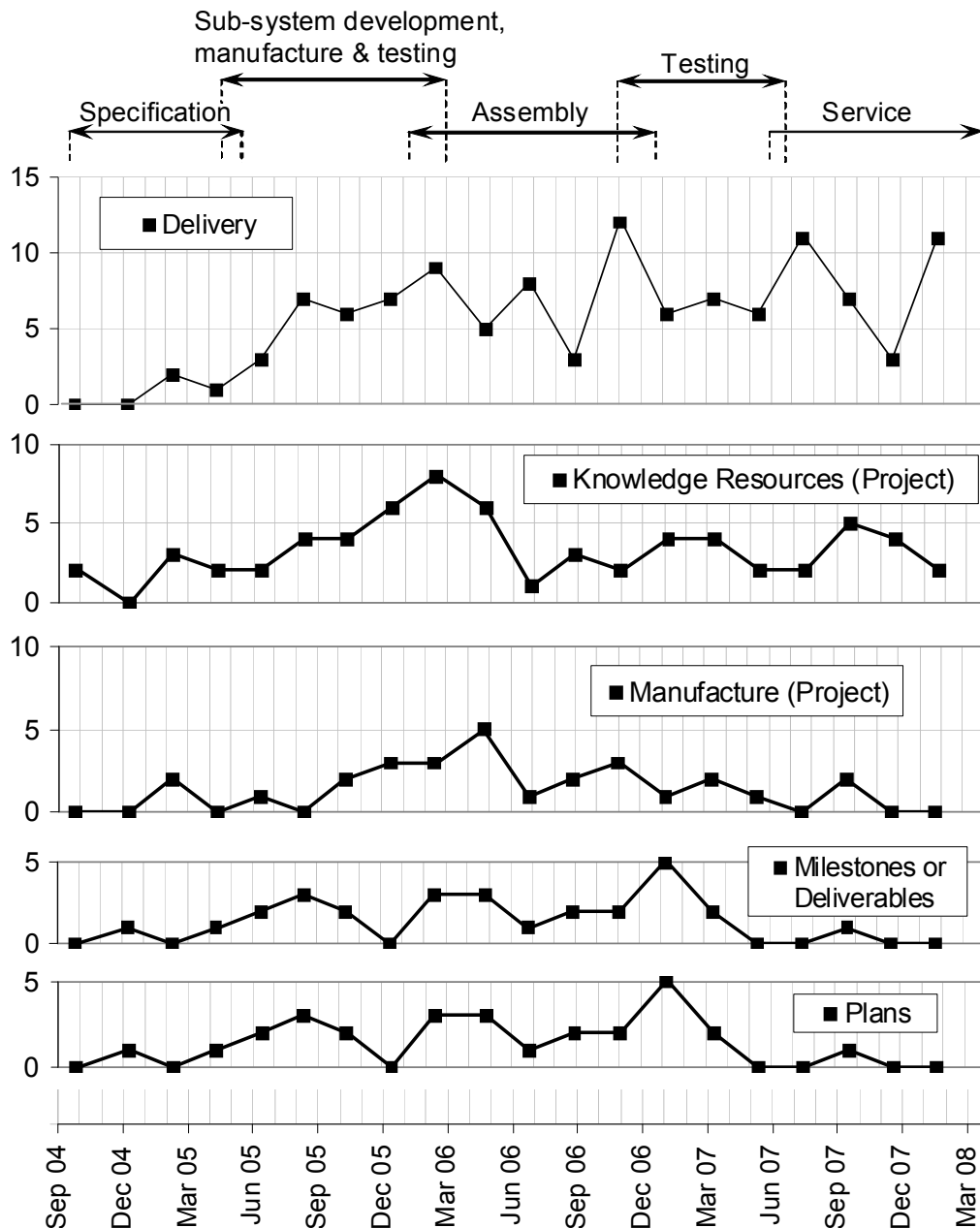


Figure 7-17: The change in frequency of emails coded to terms within the Project group over the project duration, part 1.



**Figure 7-18: The change in frequency of emails coded within the Project group over the project duration, part 2.**

Emails which are administrative (*administration*) are initially very infrequent, but rise steeply to peak just over a year into the project. After this point they drop a little but remain a fairly constant and significant proportion of emails for the remainder of the project.

The number of emails relating to *Delivery* rise as the project enters assembly stages and although with some variation stays high from then.

The number of emails discussing *Manufacturing* rise during subsystems manufacture and remain as such during assembly. It may be relevant that during this time the sub-systems for the variant products will have been produced.

### **Changes in the purpose emails were sent for**

Figure 7-19 shows how the frequency of emails sent for each purpose changed over the project duration. The proportions of emails sent for each of these purposes are seen to vary relative to on another.

*Information* transactions remain the most common purpose throughout the vast majority of the project. At one point however, April 2007 during the testing phase, there are more *Problem Solving* than either *Information* or *Management* transactions. *Management* transactions also reach their peak during this period, but also in November 2005, when *Problem Solving* is at a low. This suggests that the relationship between the occurrence of these categories may be indicative of changes in the activities and status of the project.

The changes in the use of terms which comprise the three overall purposes for sending emails are shown in Figure 7-20, Figure 7-21 and Figure 7-22.



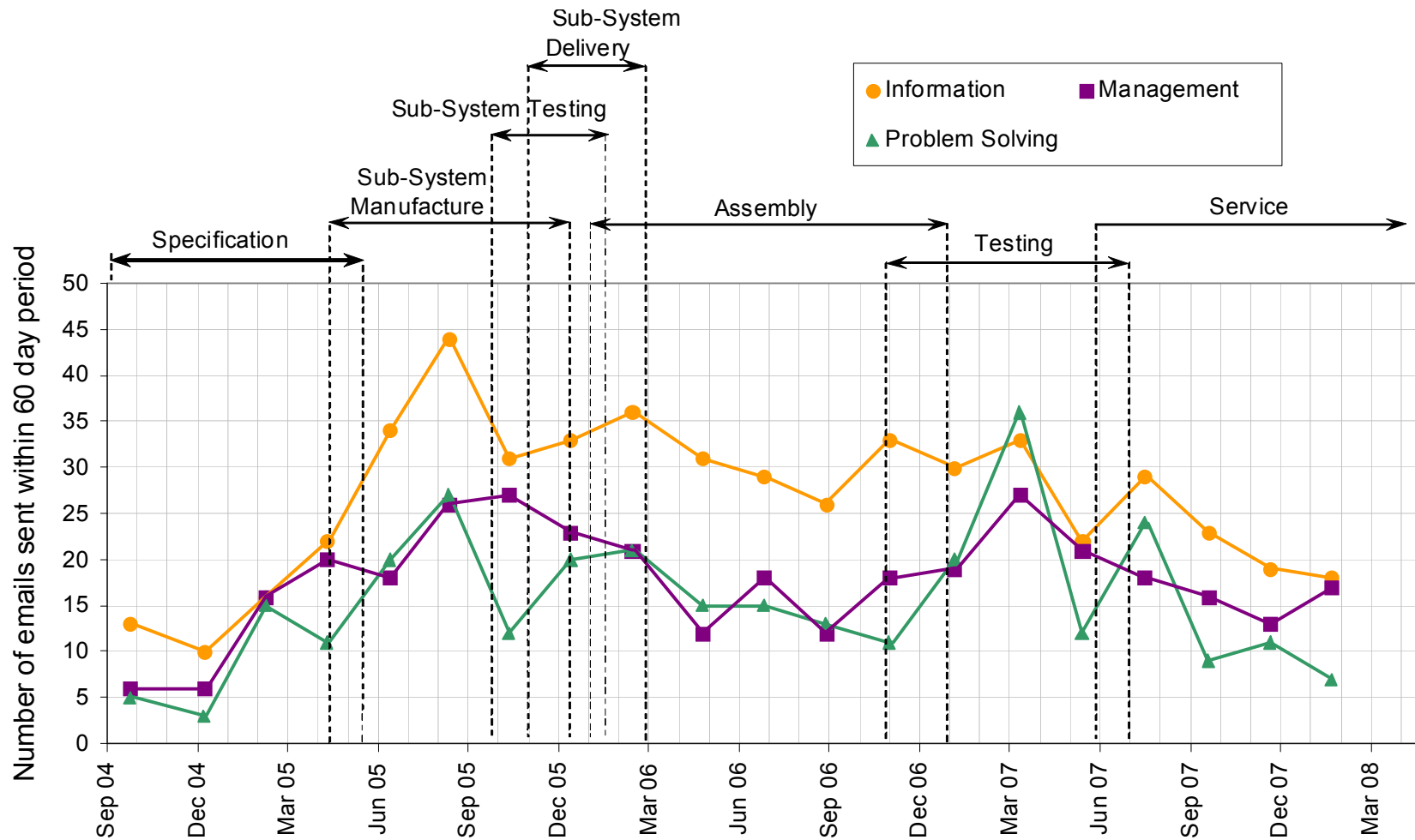
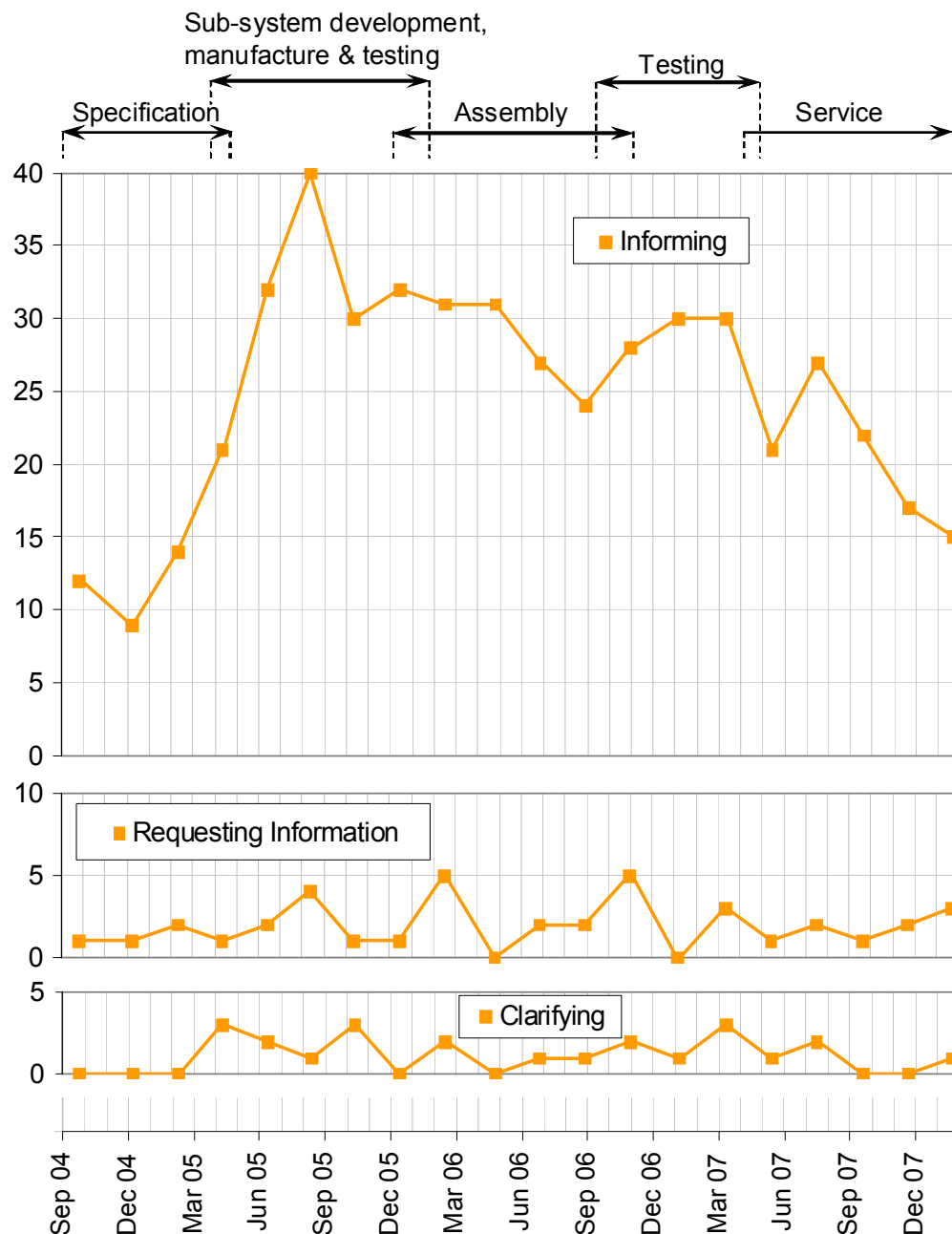
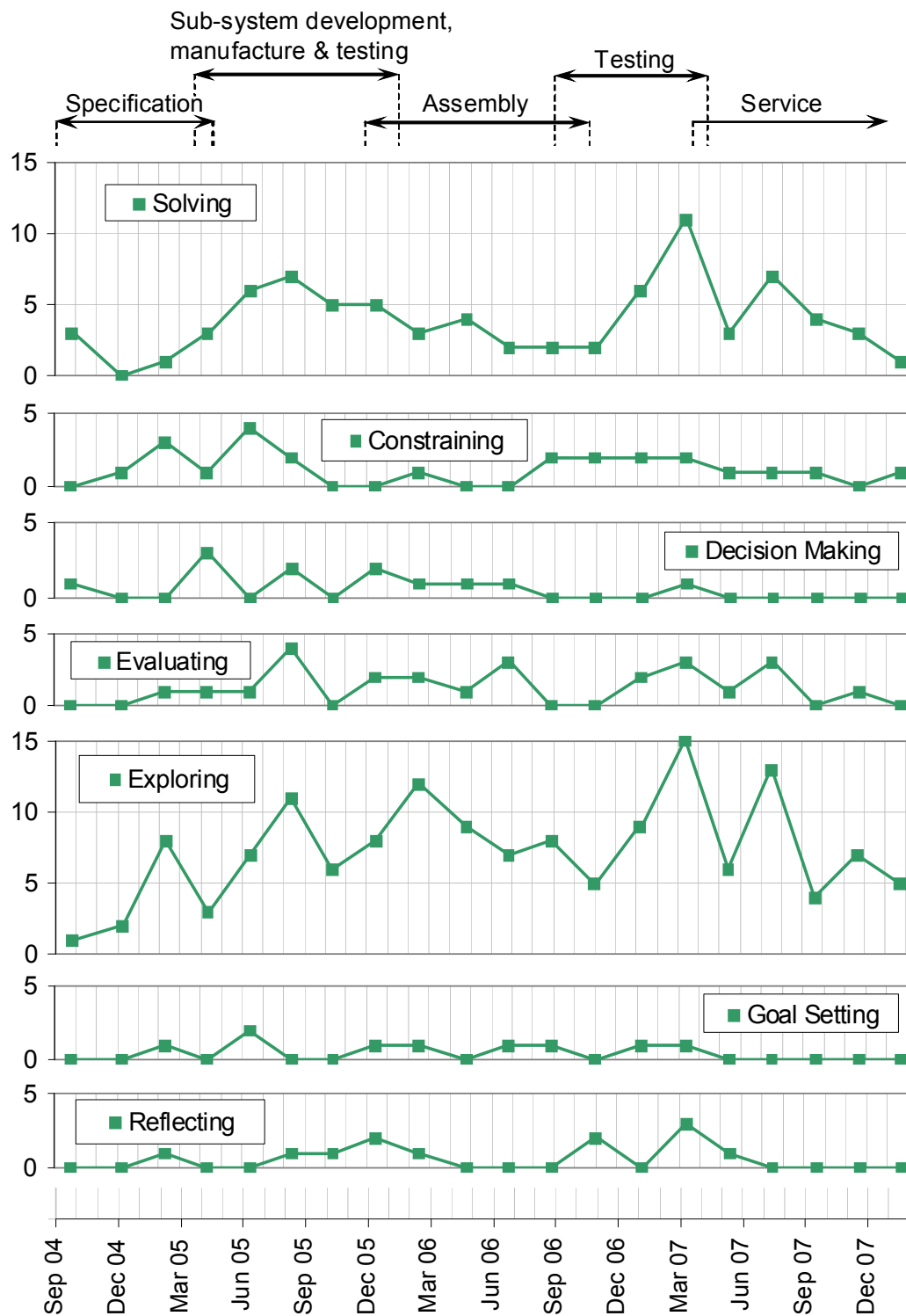


Figure 7-19: The change in the frequency of emails relating to the three purposes for sending emails over the project duration.



**Figure 7-20: The change of frequency of emails coded within the Information Transaction Group over the project duration.**

*Informing* emails rise continuously to peak one year into the project, after which their frequency tends to fall, Figure 7-20. As can be seen by comparing with Figure 7-19 this term drives most of the change in frequency of the use of the *Information* transaction group. The amount of *Requesting Information* fluctuates throughout the project, although the peaks do not necessarily align with those of *Informing*. The level of *clarifying* made remains relatively constant.



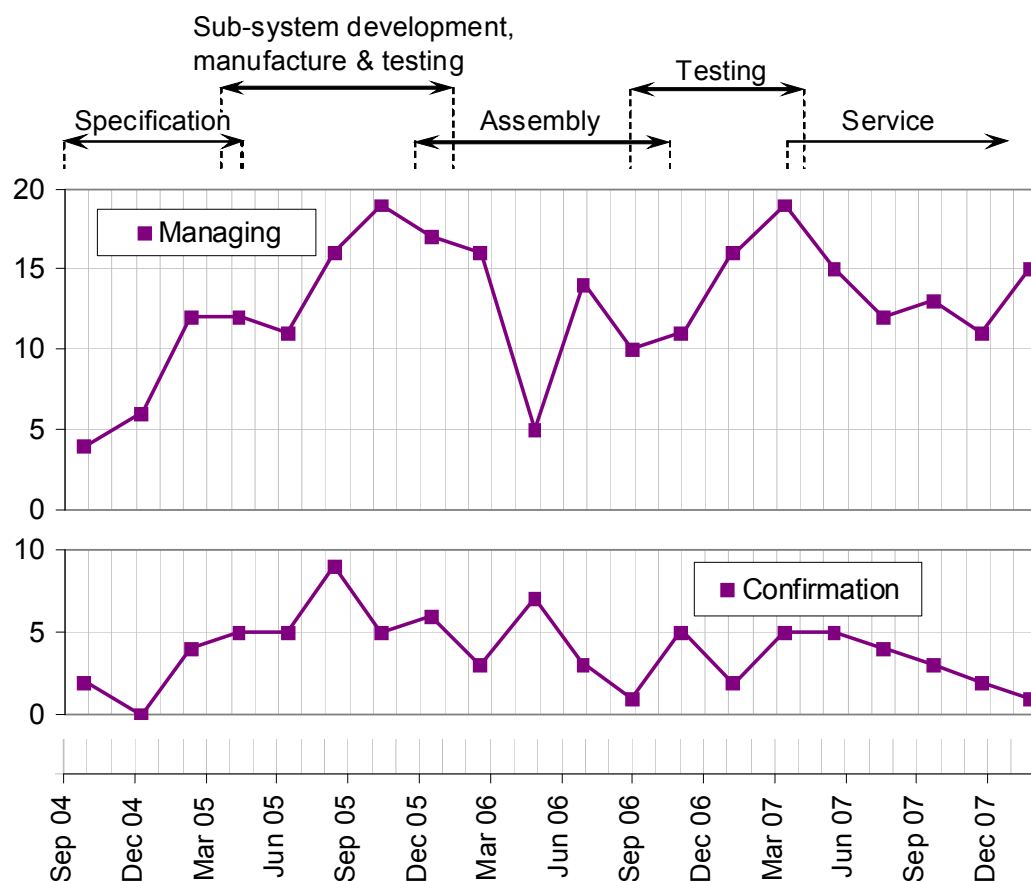
**Figure 7-21: The change in frequency of emails coded within the Problem Solving Contributions Group over the project life.**

Figure 7-21 shows changes in the different aspects which comprise the *Problem Solving* contribution group. The number of emails which discuss and propose solutions (*solving*) are relatively steady over the project.

There is however a sharp peak in activity towards the end of the testing phase from March to June 2007.

The few emails which relate to *constraining* the design or actively make decisions (*decision making*) occur more often in the early stages of the project, whilst the level of *goal setting* emails remains low throughout.

Emails which reflect upon aspects of the design or previous problem stages occur mainly during the two stages of testing. The frequency of *exploring* emails is seen to fluctuate the most, with significantly contrasting high and low points during 2007.



**Figure 7-22: The change in frequency of emails coded within the Management Transaction Group over the project duration.**

Figure 7-22 shows how the use of *Confirmation* and *Managing* emails, which make up the *Management* transaction group, changes throughout the project.

The number of *Managing* emails, like many other categories, picks up over the first year of the project. Remaining relatively frequent the one discrepancy is around April 2006, when there is a steep and temporary drop. The peaks in use of emails of this type are from October 2005 until early 2006 and again around April 2007 during the testing phase.

The frequency of emails which provide *confirmations* shows a little variation through the project life. Notably one peak coincides with the drop in *Managing* emails in April 2006. The most *confirmations* are shared by emails in August 2005.

#### **7.4.4. Results Summary**

More than half of the emails related to the *Project*, with a third discussing the *Product*. Only a very small proportion of emails discussed aspects relating to the *Company*. The emails within the *Project* group were particularly related to *time*, *delivery* and *plans*, with a quarter of all emails having an *administration* aspect.

*Information* transactions were the most common purpose for sending emails, with *Management* and *Problem Solving* each occurring in more than 30% of emails.

Two point five percent of the emails sampled, representing 400 emails in the corpus, discussed only the *product* and were sent for the sole purpose of *problem solving*. These are considered to be the emails which contain the most valuable design information.

Significant variation was shown in both the topics and purposes of emails sent by different users in the project. Those in the top managerial positions had most discussion relating to the project, with lead engineers having more discussion relating to the design. Despite such tendencies the results initially suggest that it may be difficult to predict the way in which a worker might use email based solely upon their role in the

project. For example some engineers do not appear to use their emails for *problem solving* purposes, whilst others do extensively.

The purposes for sending emails and the topics they related to were shown to change over the duration of the project. During the first year the *Product* was the main topic of emails, with the *Project* then taking over. Logistical aspects such as *delivery* tended to be discussed more at the later stages of the project, where physical integration of systems was occurring. From a design perspective, *features* of the product were considered more at the earlier stages, with *performance* considered at the later stages.

There were significant fluctuations in the proportions of *Information*, *Management* and *Problem Solving* emails over the project duration. For example during the testing phase of the project the level of *Problem Solving* emails rose dramatically.

These findings suggested that the rises and falls in topics which emails related to and particularly the purposes for which emails were sent may be indicators of project activity and success.

## 7.5. Chapter Summary

This chapter applied the approach for analysing emails to a large multinational systems engineering project, which ran over four years producing 16 000 emails. Details of the project background and company's email policy were presented. By analysing a sample of these emails and drawing some initial findings, this addressed **Objective 4** of the research: *To apply the proposed approach to recently completed industrial projects and to explore the role that email plays over the project life.*

Specific details concerning the application of the coding scheme and approach to this case study were then described. This included how the emails were sampled and the level to which different elements of the coding scheme were applied. The other data collection through interviews and reviewing project documentation were also described. The metrics for reliability and the validity of the coding were then reported.

Three sections of results were presented and discussed. The first presented the frequency with which the different codes appeared in the email corpus, this also explored combinations of coding. The second section presented how the use of codes varied between emails sent by different users. The final section showed how the coding of emails changed over the duration of the project.

## 8.0 Findings: Discussion and Analysis of Case Studies

This chapter analyses evidence, in the form of results from the pilot and major case studies (Chapters 6 and 7 respectively) and literature reviewed previously, to arrive at prototypical findings regarding how email is used in an engineering design context. This addresses research **Objective 5**: *To elicit general observations regarding the role and use of emails in engineering projects.*

The format of this chapter is to state each finding, followed by a discussion of the evidence to support this. The findings are arranged in the order of the three dimensions analysed in the previous results chapters, these are; frequencies of coding elements, variation with users and variation with time. The findings in this chapter are based on results from case studies.

The first section provides a summary comparison of the pilot study and major case study and explains what is meant by project characteristics, as this is relevant to the following sections of findings. The next section presents general findings relating to the characteristics of emails. The following section presents the influences of people, locations, job roles and the company on these. The fourth section explores how the topics of email change over time. The fifth section then explores how purposes of email change with time and more specifically how these represent signatures of design activity.

### 8.1. Comparison of Case Studies

The pilot and major case studies vary greatly in their characteristics. The background to the pilot and major case study projects was presented in the early parts of Chapters 6 and 7 respectively. The differences are summarised here in Table 8-1. Herein the phrase ‘project characteristics’ is often used; it refers to such facets as identified in Table 8-1. These are



the domain, type of design, duration, budget, number of participants, extent of geographical distribution and whether the project is interdisciplinary. Throughout the analysis, project characteristics refer to the values held for each of these.

**Table 8-1: A Comparison of Characteristics of the Projects from the Pilot and Major Case Studies.**

	<b>Pilot Case Study</b>	<b>Major Case Study</b>
Parties	Intra-company (Single company, two sites)	Inter-company (Multisite, multi partner)
Domain	Software Engineering	Mechanical Engineering Software Engineering Systems Engineering
Disciplines	Single Disciplinary	Interdisciplinary
Type of Design	Variant Design	Adaptive Design
Duration	2 year duration	4 year duration
Distribution	2 locations	3 principal locations. Numerous locations for other suppliers
Participants	15 core participants	30 core participants
	63 email senders	650 email senders
Budget	£100K's	£M's

## **8.2. Characteristics of Emails**

The first set of results in each of the case studies showed the overall proportions and combinations of email purposes and topics. The following findings derived from these results (from both case studies) and existing literature describe general characteristics and tendencies in the use of email in engineering designs. A comparison of the two case studies also leads to findings as to how the characteristics of a project or team influence email use.

### **Finding 1. The predominant purpose of emails is to transact information**

The results of the major case study show that the predominant purpose of sending emails is to achieve an information transaction. Thirty eight percent of emails were sent solely for this purpose, and 73% for this purpose alongside another. Thirty nine percent and 86% were found respectively in the pilot case study, strengthening the assertion that information sharing is the principal purpose of email use in engineering design contexts. This is reflective of findings in the existing literature regarding email use in general; in particular the work of O'Kane et al., (2007) who found that a principle purpose of emails was the transferring of information, passing views, reports or ideas from one person to another.

Most information transactions were informing (60% of all emails), rather than requesting information or clarifying. This shows that most emails particularly support the 'handover' communication scenario defined by Eckert et al., (2005) in which one person passes their work on to another, not requiring any feedback.

### **Finding 2. Emails are used for directing activity and management**

The results show that emails are used for directing activity and management purposes. Nineteen percent of emails in the major case study were sent simply to achieve a management transaction, with 45% of emails including a management transaction with a further purpose. Similar frequencies were found in the pilot case study, 16% and 34% respectively. A significant number of emails relate specifically to project management activity, of this, 34% contained aspect of both a management purpose and a project topic.

Existing literature does not suggest to what extent emails are used as a management tool, although several authors (e.g. Panteli, 2002) recognise this as a role. What is clear from the case studies, and supporting

interviews, is that emails are used not only by managers to direct activity, but by engineers to request actions of their managers and peers.

**Finding 3. Emails are used for exploratory discussion and solution development, both of which are evidence of problem solving.**

In both of the case studies exploratory discussion (*exploring*) took place in around 20% of emails and problem *solving* in around 10%. Both of these creative elements provide information which can later be used in the decision making process. The presence of these elements would support the view that email is used for decision making purposes as contended by Lusk (2006). However in this study, only the development of solutions is found to be discussed and not the final decision making stage.

**Finding 4. Few decisions are clearly made and expressed in emails.**

Although there is evidence of exploratory discussion and solution development in emails, progression from this into decision making is less well evidenced. In both case studies the frequency of *decision making* was very low, occurring in less than 1% of emails. It was recognised that decisions made in emails may not have been clearly manifested as such, because the preceding emails in the conversation were not reviewed. In these cases they may have been interpreted as instructions (*management transactions*) or information sharing.

Further to this; where a decision is made in a face-to-face conversation, but confirmed in an email, the coding of *confirming* rather than *decision making* could be applied. This will depend on the wording of the email and whether it is sufficiently clear a decision is being made. For example an email stating “*I can confirm that we have selected, and are ready to proceed with ordering the roller bearings*” would be coded as a decision making and confirming. An email stating “*please accept this email as confirmation for our roller bearing order*” does not contain sufficient

evidence to be coded as a decision making. (Refer back to Figure 5-4 for the full example text).

The coding scheme captures decisions explicitly made in emails. It may however fail to identify decisions which are presented as confirmations. It is noted that confirming is far more frequent, occurring in 10% of emails in the major case study. Based on the experiences of coders, it is likely that only a small proportion of these will contain decisions which were not clearly identified. On this basis it is found that few decisions are clearly made and expressed in emails.

### **Finding 5. Design Work is Facilitated via Email**

Emails that transacted information about the product or that managed activities relating to the product, as identified by the coding scheme, were regarded as facilitating design work. They provided information or direction necessary for the design process without actually explicitly requesting the design work.

Evidence for how frequently this occurred comes from Table 7-5, which showed combinations of email purposes and topics of discussion. A significant number of emails (7.6%) solely transacted *information* about the *product*, and 3.2% *management* and the *product*. A far greater proportion, 12.9% were seen solely to facilitate design work, containing only *product* discussion with an *information* purpose or *product* discussion with a *management* purpose. When emails which also had an element of *project* topic matter are included, this figure rises to 23.1%.

Emails of this nature support begin to support the ‘interface negotiation’ communication scenario defined by Eckert et al., (2005). Participants are sharing information which can make elements of the design consistent and discuss different possible ways of achieving this. These emails do not however show the consensus and decisions reached to complete the scenario.

### **Finding 6. Design Work Takes Place Within Emails**

It was argued in Chapter 2, that engineering design can be regarded as a problem solving process and hence emails which include *problem solving* around the topic of the *product* are seen to represent design work. Those with the only topic of the *product* and only sent for the purpose of *problem solving* explicitly show design work taking place in the email. This implies that such design work is not being recorded elsewhere. Other emails containing the *product* as a topic with a *problem solving contribution* purpose which also contained other topics and purposes will have facilitated designing, even if the design work is not actively taking place in the email. Emails which contained problem solving but do not have the product as a topic are not seen to contain or facilitate design work but instead relate to project management.

In the major case study 2.5% of emails had the *product* as their only topic and *problem solving* as their only purpose. These are the emails in which designing is taking place (rather than simply being facilitated). Whilst 2.5% may appear a low figure, in this corpus of 16 000 emails this equates to 400 emails which in turn equates to around 8 000 lines of information.

In addition to this 2.5% it is suggested that an element of design work is likely to take place as part of many of the 14.6% of emails which included the *product* in the topic matter and *problem solving* as a purpose but may also have contained *project*, *management* and *informing* activities. Recall that for the purpose of this study only emails that contain solely *product* and *problem solving* are considered to be design work. This is because the problem solving can be explicitly associated with the product.

Although less prevalent than *information* transactions it is clear that emails are sometimes used for actively doing design work. This again ties with the findings of O'Kane et al., (2007) that the generation of ideas in emails is subservient to information sharing.

### **Finding 7. The way in which emails are expressed will be influenced by the relationships between team members**

Eckert et al., (2005) describe communication processes as socially orientated, and it could be intimated that the status of users' relationships with one another will affect the way in which their emails are communicated.

In the pilot case study, the participants were part of an established community outside of the email environment. They were closely located geographically and had worked with each other before. This contrasts with the major case study, where more participants were distributed and had not known each other previously; and might not even have met.

When considering *how* email content was expressed in these projects of contrasting social structure (Figure 6-8, Figure 7-10) key differences can be seen.

In the pilot case study (Figure 6-8) with the closer community there are three times as many expressions of *solidarity*. There is also the presence of *tension release* which is not demonstrated in the major case study. It may be that expressions of *solidarity* and *tension release* are not necessary for the understanding of an email, but act to strengthen the bonds within a team. The group in the pilot case study also asked for *orientation* and *suggestion* from their team members twice as often, although this may be a result of the project characteristics.

In both case studies the expression of *tension* is equally well represented. This suggests that close knit communities are likely to share both positive and negative emotions, whilst less well connected groups are good only at sharing negative emotions when things go wrong.

### **Finding 8. The characteristics of a project will affect the subject matter of emails**

In the pilot case study the majority of emails related to the *company* category, whereas in the major case study the majority of emails related to the *project*, and very few to the company *category*.

It is contended that the contrasting characteristics of the projects resulted in this variation. The pilot case study was an internal project, which explains the influence of company issues. The major case study was heavily connected with a number of external parties, and it is logical that emails relating to the *company* were less represented. This supports the argument that the characteristics of a project will affect the content of emails.

### **Finding 9. The characteristics of a project have a limited affect on the purposes for sending emails**

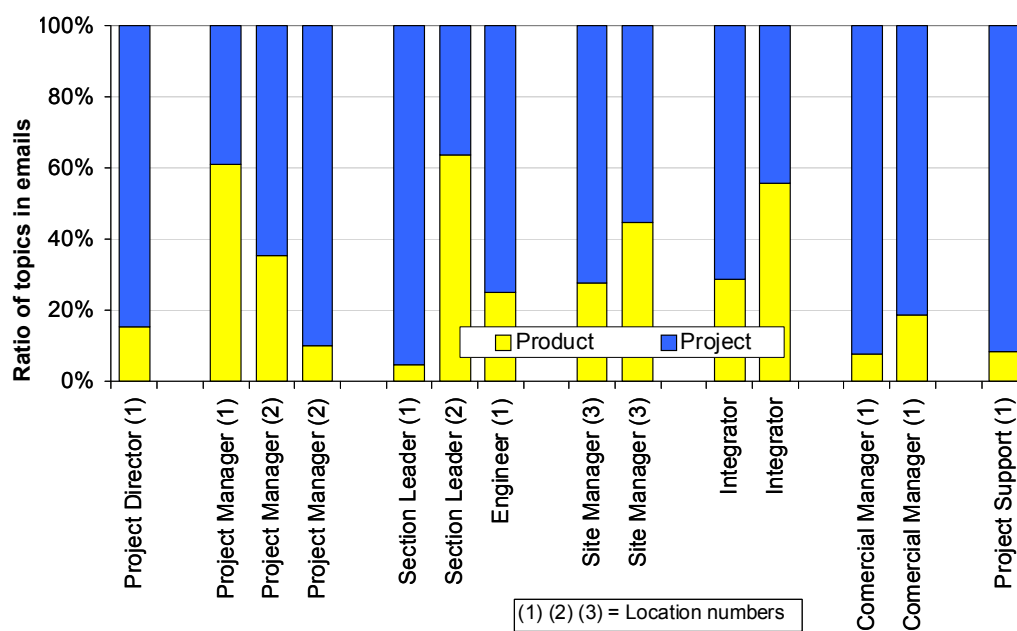
Although the two case studies had different characteristics, the purposes for sending emails were very similar between the case studies, as can be seen by comparing Figure 6-6 and Figure 7-8. For example between the major and pilot studies respectively: 38% and 39% have only an *information* purpose; similarly 6% and 7% have only a *problem solving* purpose. Finally 15% and 16% have both an *information* and *management* purpose. The minor exception to this tendency is *management* only emails, which occur four times as often in the major case study.

The overall similarity is surprising because the senders are not consistent between the email corpuses, and in particular there are many users from different companies and hence email cultures in the major case study.

This would suggest that the purposes for sending emails, in a engineering design context at least, could be relatively consistent between corpuses from different projects and companies.

### 8.3. Influences of Person, Location, Role and Company on Email Use

The second dimension of analysis showed how email use, topic and purpose varied between different users. Given the different roles, location and parties the users represent, the following findings are drawn on how these factors influence email use. Again, results from both pilot and major studies were used to draw these findings. To aid the reader, some figures from Chapter 7 are now reprinted.



**Figure 8-1: The proportion of Product and Project coded emails sent by key users in the industrial partner company. (Reprint of Figure 7-11).**



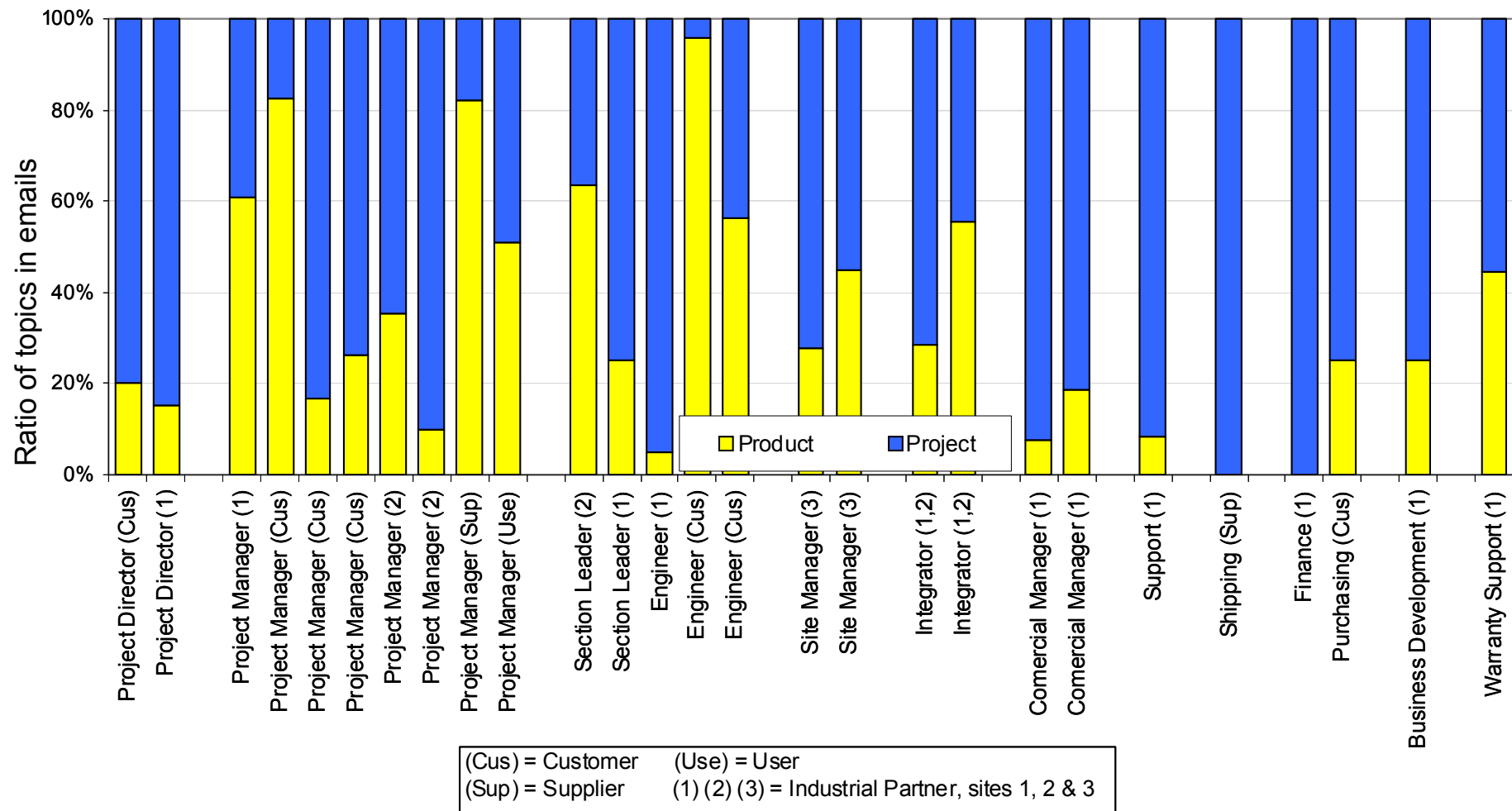
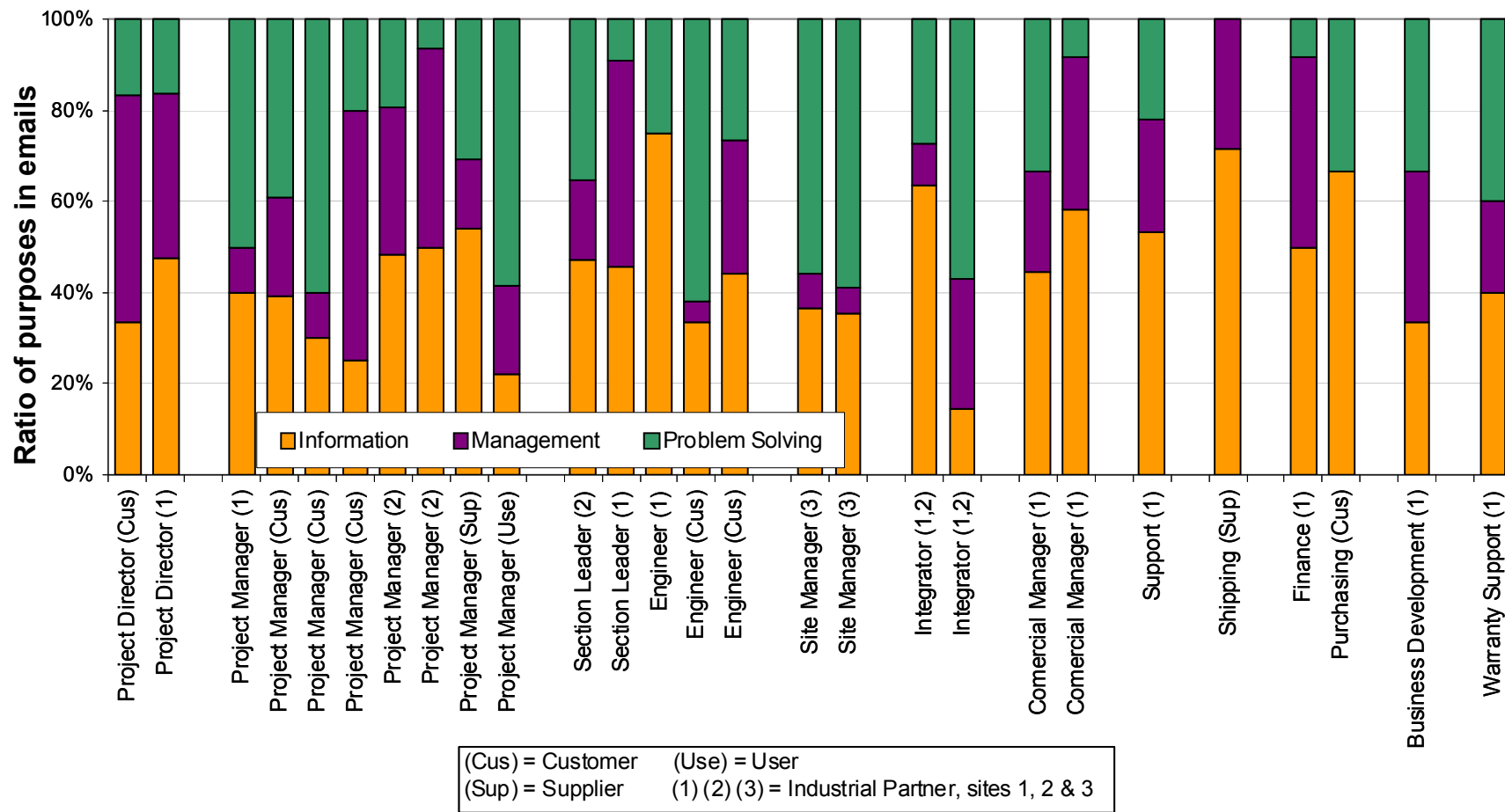
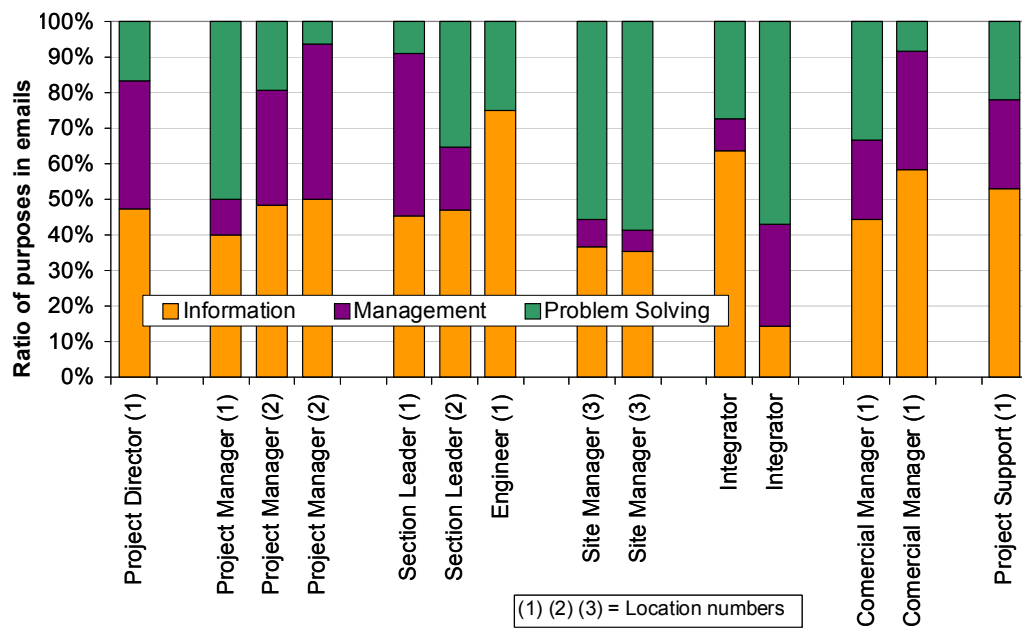


Figure 8-2: The proportion of Product and Project coded emails sent by key users from major parties in the project (Reprint of Figure 7-12).



**Figure 8-3: The proportion of Information Transactions, Management Transactions and Problem Solving Contributions coded in emails sent by key users from major parties in the project (Reprint of Figure 7-14).**



**Figure 8-4: The proportion of Information, Management and Problem Solving purposes coded in emails sent by key users in the industrial partner company (Reprint of Figure 7-13).**

### **Finding 10. Inter-Company emails often discuss the product more than Intra-Company emails.**

The major case study corpus analysed included a mixture of intra-company and inter-company emails sent by users from the industrial partner company. Only inter-company emails sent by users from other project partners were captured.

The correspondence of user's in engineering roles from these external parties tended to be more heavily dominated by the *product* than the emails of the industrial partner's users, Figure 8-2. For example one customer engineer sends correspondence which is approximately 95% *product* related compared to 55% for his counterpart from the industrial partner. This suggests that inter-company emails are more likely to relate to the product and its design.

One interviewee, an engineer, said that he was unlikely to send emails discussing the product to his colleagues in the company as he would

interact with them *face-to-face*, but would discuss product detail in an email with the customer. Location can therefore be a major influence on the difference between inter-company and intra-company emails.

**Finding 11. The purposes for sending emails are consistent between inter-company and intra-company messages.**

Given the link between inter-company emails and prevalence of *product* discussion; it might be expected that a correlation with an email sending purpose would also exist.

This was shown not to be the case with the varied purposes for email sending bearing no correlation to the parties involved, Figure 8-3. In most cases the purposes of engineers' and managers' emails were between 40% to 50% *information* transactions regardless of which party they represented. The reasons for email communication at inter-company and intra-company levels are therefore the same. This differs from finding 10, where the type of information discussed via email (i.e. product information) is influenced by which parties the email is between.

**Finding 12. Location of parties affects the subject matter of emails**

As has been discussed a possible reason for the observed characteristics of inter-company emails could be the distributed location of the parties. Further evidence of location influencing the subject matter of emails is found in the emails of the site managers and integrators, Figure 8-1.

All of these users were employees of the industrial partner, whose email communications were largely intra-company. However, the site managers and integrators mainly worked from distributed locations to those they communicated with. Their emails were seen to contain a larger proportion of *product* information than many of their peers in the main location. For distributed roles this was 30% to 55% compared to 5% to 25% for most of those at the main location (Figure 8-1). An exception to this was a section leader whose role involved communicating with those at the assembly

site. Although the observed trend may be partly attributed to the job roles of these users, the interview with the project director suggested that they should still be discussing project issues.

This suggests that product related discussion is heavily influenced by location, and where it is possible to discuss the product *face-to-face* users find this to be preferable.

### **Finding 13. Personal preferences affect how email is used**

The existing literature suggests that user's choices regarding when and how to use email are highly idiosyncratic (Wilson, 2002). The interviewees from the case study generally expressed differences in attitude towards when and how they used email, supporting these existing findings. Such differences relate to familiarity with technology and the ability to use email over alternative mediums, and personal opinions as to which mediums are more suited to different tasks.

Figure 8-2 and Figure 8-3 present the topic and purpose of emails and show significant variations in email use between the many project participants. The user labelled 'Project Manager (1)' is of particular interest as he was also interviewed. When compared with his peers who had similar roles (labelled 'Project Manager (2)') this user's email involved approximately 80% fewer *management* transactions, Figure 8-4. When interviewed the user expressed that he would always prefer to communicate *face-to-face* whenever possible and that emails would then only be sent if a written record of an instruction was needed. It was acknowledged that this person tended to facilitate more of his management responsibilities *face-to-face* than his peers, as the results reflect. The user also tended to discuss the Product more than most managers, Figure 8-1. This is seen to be reflective of the user's attitudes, who during his interview, expressed that he considered himself to be one of the engineers within the team he managed.

#### **Finding 14. A user's job role has an affect on the topics and purposes of their email**

From the interviews it was clear that users in different roles used email in (subtly) different ways to support their work. The results show some tendencies for both topics and purposes of emails used by different job roles.

##### ***Topic***

Common wisdom would suggest that the subject matter of a user's emails would be related to the job role they fulfil. This is clearly demonstrated in the results of the major case study, Figure 8-1.

The users who have non-engineering roles, such as the commercial manager and project support workers, send far more *project* emails than the users in engineering roles, whose correspondence involves more *product* related content. Their emails never rise above 20% for *product* related content compared with 20% to 65% for most other roles.

A more specific example is seen with the two Project Directors who, although from different companies, send a very similar ratio of *project* to *product* emails, 20% to 80% for one, 15% to 75% for the other, (Figure 8-2). Other pairs of users, such as the commercial managers also show strong similarity with one another.

There are however a number of other groups of users with similar job roles whose email topics are not consistent with one another. This may in part reflect the limitations of a job title for describing adequately what a participant does. Whereas the Project Director and Commercial Manager roles are well defined, there is a large scope between the different Engineering and Project Manager roles; for example there are software engineers, systems engineers and mechanical engineers.

### ***Purpose***

A number of general trends are seen in the purposes for sending email, which correlate well with users' roles.

The Project Director for the industrial partner uses email for *management* transactions on 40% of occasions, reflecting his role in the project. The counterpart Director for the customer shows a similar profile; sending 50% of emails for this purpose (Figure 8-3). Both users send approximately 17% of their emails for *problem solving* purposes.

Further evidence for a link between role and purpose of emails is that the users who engage in most *problem solving* via email, tend to be in engineering roles.

The two site managers are seen to use email for very similar purposes, Figure 8-3. Both send 35% of emails for *information* purposes, between 7% and 10% for *management* and between 55% and 58% for *problem solving*. Notably, because of their remote location and difference in time zone from the team members they most often communicate with, they are very much constrained to using email. It is therefore argued that job role does impact on purposes for email sending, but that other factors are likely to dominate over this, such as personal preferences of users and their geographic location.

## **Finding 15. Combinations of factors affect the topics and purposes of email**

The previous sections have considered how a number of different, sometimes interrelated aspects, affect the topics and purposes of emails.

**Table 8-2: Factors affecting the use of email**

<b>Factor</b>	<b>Topic</b>	<b>Purpose</b>
Inter vs Intra Company Email	✓	✗
Location of Users	✓	✗
Job Role of Users	✓	✓
Personal Preferences	✓	✓
Project Characteristics (previous sections 8.1 & 8.2)	✓	✗
✓ = some correlation      ✗ = little correlation		

From the case studies considered, it is very hard to isolate any one of these factors to explore in detail. What can be said with a good degree of certainty is that a combination of several key factors will affect email use. As a summary of these, based on the findings discussed, Table 8-2 shows where there is strong evidence of a link between the factors considered and the topic or purpose of an email.

### **8.4. Changes in Email Topics with Time**

The final analysis of results from the case studies showed how purposes and topics of emails changed over the project duration. In this section the findings derived from this explore the ways in which topics reflect activity and progress in a project.



### **Finding 16. The topics of emails reflect the stages and activities of a design project**

It seems intuitive that project activities should be discussed in emails, and hence that the subject matter of emails will reflect the activities and design stage of the project.

In both case studies the relative frequencies of *product*, *project* and *company* emails varied throughout the project duration, along with the frequency of email exchange. It is argued that the relative frequency with which topics occur within emails is reflective of the priorities of the project at any given time. Examples evidencing this are now presented.

In the major case study, during the initial specification stage, more emails related to the *product* than the *project*, Figure 7-15. However, once the manufacturing phase was established, more of the emails related to the *project*. This aligns with the understanding of engineering projects discussed in Chapter 2. During the specification stage of a project there is a focus on the product and determining what it needs to do. In the later stages, although design work progresses, the emphasis shifts to aspects of project control, such as time, budgets and resources. This is reflected in the detailed breakdown of email topics from the major case study, where *time* and *delivery* aspects increased, Figure 7-17.

Another example of the link between project activities and their discussion in email topics, is revealed by comparing product and project costs. The results show *product costs* were discussed more in the early stages while *project costs* were discussed later on, Figure 7-16 and Figure 7-17 respectively. This reflects the model of costs being committed in the early product design stages, but incurred later on in the project, as discussed in Chapter 2.

A final example is the increase in *product* emails that was seen in February to April 2007 (Figure 7-15). From the supporting project

documents and interviews it is known that this period involved the testing of the product, and highlighted some significant issues in its performance. The steep change in *product* type email activity is a clear reflection of this, along with the increase of references to *performance* in emails at this time, Figure 7-16.

These examples in particular, provide evidence that the topic matter of emails is reflective of tasks and activities in the project.

## **8.5. Signatures of Design Activity and Project Progress**

In this section the purposes for email sending and how these vary with time are explored. In particular signatures of email are used to identify project progress. Figures from Chapter 7 are reprinted to aid the reader.

### **Finding 17. Email purposes can be signatures of design activity and project progress**

The results in 7.4.3 identify a number of purposes for sending emails (*Information Transactions, Management Transactions and Problem Solving Contributions*) which as for email topics, vary in their frequency of use throughout the project. These are shown in Figure 8-5 and Figure 8-6. It was found that these trends strongly reflect project activity and progress and that their relative signatures implicitly represent design activity.

It was possible to map email use against the phases which took place in the project and in particular whether the number of emails relating to *information and management transactions* and *problem solving contributions* was rising or falling. This was contextualised with respect to the project plans and progress, the latter of which is elicited from the interviews.

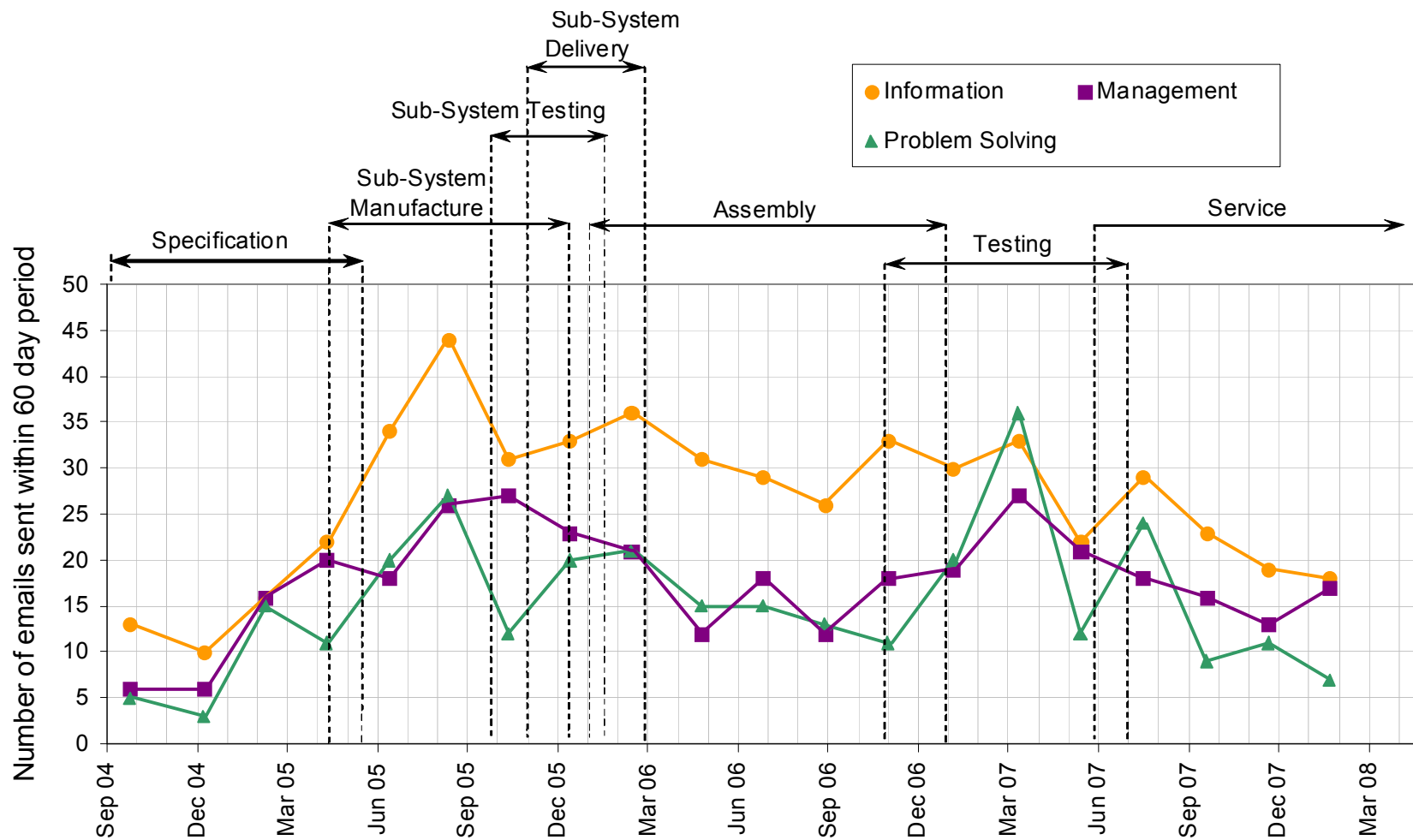


Figure 8-5: Change in the frequency of emails relating to the three purposes for sending emails over the project duration (Reprint of Figure 7-19).

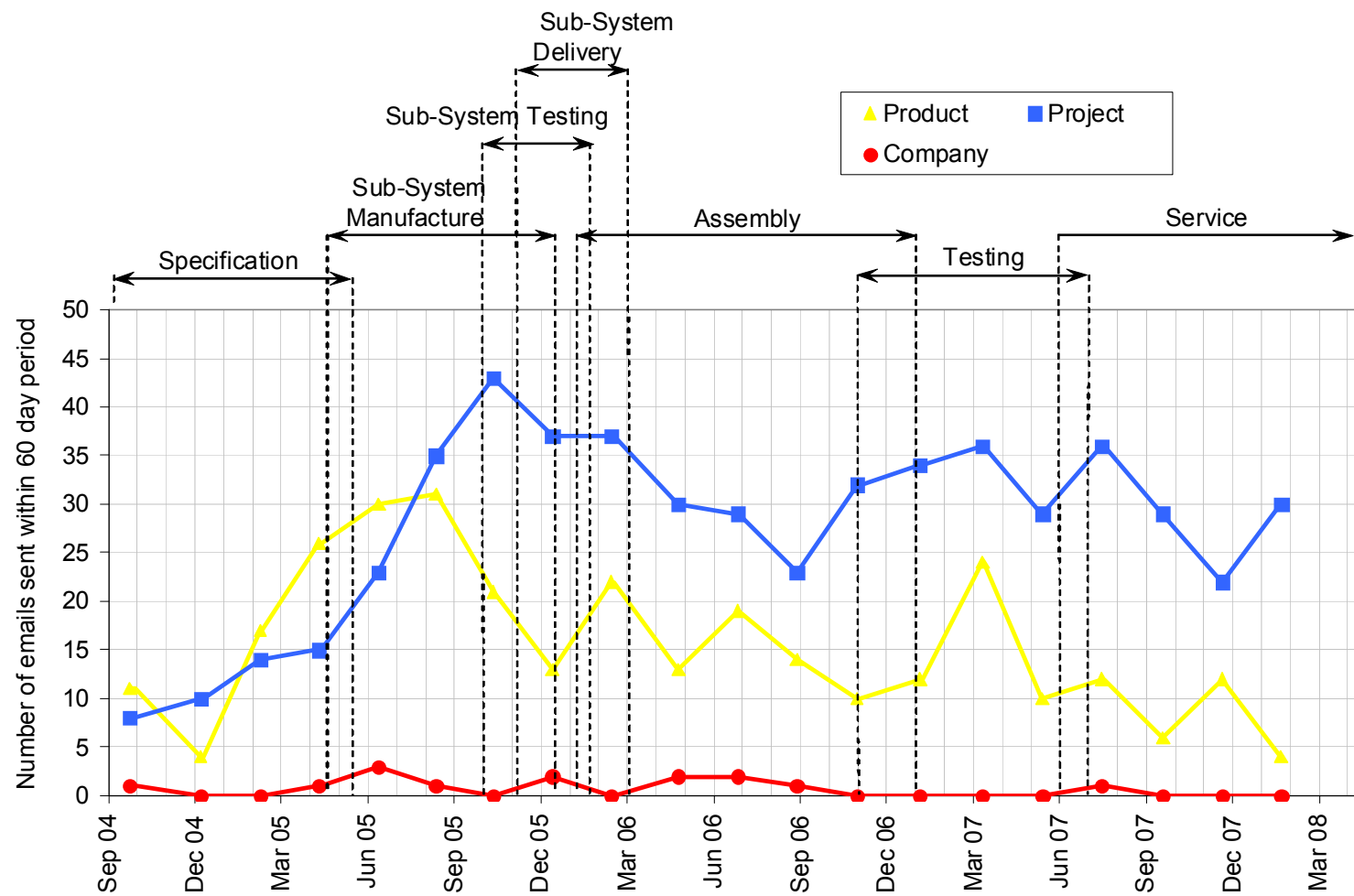


Figure 8-6: The change in the frequency of emails relating to the three topic groups over the project duration (Reprint of Figure 7-15).

The possible permutations of the rise and fall of e-mail traffic related to *information, management, and problem solving* are shown in Table 8-3, along with hypothesised interpretations with respect to engineering projects. A rise is defined as an increasing level of email sending and a fall a decreasing level of email sending (in relation to any given purpose).

**Table 8-3: Permutations of rises and falls in information, management and problem solving purposes and interpretations, drawn from Figure 8-5.**

	Management	Information	Problem Solving	Interpretation	Example Date
1	Rise	Rise	Rise	Pressure point, with real co-ordinated work achieved.	6/05 – 8/05, 1/07 – 3/07
2	Rise	Rise	Fall	Work required but progress ineffective, further information may be required.	9/06 – 11/06, 2/05 – 4/05
3	Fall	Rise	Rise	Progress is made with little management direction.	10/05 – 2/06, 5/07 – 6/07
4	Fall	Fall	Fall	Steady mode of working	3/06 – 9/06
5	Rise	Fall	Fall	Management impetus to increase work	5/06 – 7/06
6	Rise	Fall	Rise	Information required for work available	11/06 – 1/07
7	Fall	Fall	Rise	Direction clear and information sufficient	10/07 – 11/07
8	Fall	Rise	Fall	Management intervention and improved problem solving required to improve information flow	Unobserved

All three purposes rose significantly during two phases of the project; June 2005 until August 2005 and January 2007 until March 2007 (Figure 8-5). At the same time, the patterns in Figure 8-6 reveal that the *product* related e-mails also increased. Both of these periods have also been identified as key pressure points in the project, especially the second phase, which shows the steepest increase. The simultaneous and

sudden rise of all three purposes along with increasing rates of *product* and *project* email was thus a predictive signature of heightening work activity to deal with the problems that arose during the transition from design engineering (final specifications) to manufacturing.

The trend of decreasing *problem solving* contributions in the email is associated with periods of work progress or completed work. During the period between September 2006 and November 2006 and to some extent February 2005 until April 2005, both *information* and *management* discussions rose, whilst *problem solving* did not. During this period of assembly and test set-up, there was a high level of discussion relative to sharing information and managing the work, but the lack of *problem solving* suggests that little work was being achieved. It is possible that the team could not complete their tasks or did not feel sufficiently informed to carry out their tasks. Such a scenario may be an indicator that a pressure point, described previously, is about to occur.

One of the most interesting trends is when the level of *information* and *problem solving* contributions rise, yet *management* falls. This period is identified as one of effective work where individuals are communicating as necessary with each other to complete their work with minimal administrative intervention. This occurred from October 2005 until February 2006 and again briefly in May and June 2007.

When all three purposes fall this is a predictive signature that the project is progressing steadily and without concern. The explanation for this trend, seen generally between March 2006 and September 2006, is supported by interview findings. This was demonstrated in the smoothest phase of the project, assembly, when there was a general decrease in the rates of all e-mail purposes and all topic groups. At a brief point between May to July 2006, the level of *management* increased even while *information* transactions and *problem solving* were decreasing. The secondary data indicates that some management intervention was necessary at this point to direct the project. Nonetheless, a general

decrease in *problem solving* was evident at the end of each phase. This trend is to be expected since the need to share information of what work has been completed and *management* transactions to verify milestones and contractual obligations would happen at the end of each project phase. Of concern would be if *problem solving* contributions increased toward the end of a scheduled phase, indicating the potential of missing a deadline.

Periods of brief increases in *management* related transactions were encountered throughout the project. Between November 2006 and January 2007, *information* transactions fell whilst *management* and *problem solving* rose. Management intervention was needed throughout the project to ensure that contractual obligations were being met, even when all necessary information for work to proceed was present.

Likewise, from October to November 2007, the amount of *problem solving* increased, without further *management* or *information* transactions. At this point, the product was entering service delivery and the e-mail traffic confirms that the product met contractual obligations with only minor 'teething' problems. The direction of work was clear and there was sufficient information.

At no point did the level of *information* increase whilst *problem solving* or *management* decreased. This is unsurprising as many *management* and *problem solving* emails also transacted *information*. However if this were the case it should indicate that both management intervention and improved problem solving from the team are required in order to improve information flow.

In summary, based on the observed trends and explanations it is argued that the purposes of emails can be signatures of project progress and design activity.

### **Finding 18. Email use is dynamic**

As an overarching finding, it has been shown that email is used dynamically over a project. In the major case study, the number of emails sent for a certain purpose or relating to a specific topic was often double or triple in some months when compared to others, Figure 8-5 and Figure 8-6. Although the number of hours worked each month varied over the project (due to increases in staff, overtime and holidays) it never doubled or indeed tripled. The two and three fold increases in email use can only therefore be attributed to a requirement to communicate for different purposes determined by the progress and stage of the project. This finding concurs with theoretical expectations, and links back to the perspective of Eckert et al., (2005) regarding communication as dynamic and temporal; contextualised by time frames and relating to other events.

## **8.6. Summary**

This chapter discussed findings which were drawn from the results of the case studies and the literature review. In doing so it addressed research

**Objective 5:** *To elicit general observations regarding the role and use of emails in engineering projects.*

It was identified that the main purpose of emails is to enable information transactions. In the case studies 73% and 86% of emails had this purpose alongside either a management or problem solving purpose. Emails are also commonly used for directing activity and management with nearly 20% of emails in the major case study sent only for this purpose.

The earlier stages of problem solving processes, such as exploratory discussion and solution development were found to be supported by emails; taking place in 20% and 10% respectively. However decisions arising from these were only found to be explicitly stated in less than 1% of emails. It was identified that between 13% and 23% of emails were



used to enable and facilitate design work taking place elsewhere. However, only in around 2.5% of emails did design work take place in the message content itself.

It was found that the way in which emails are expressed is influenced by the relationships between team members. As intuition would suggest, the characteristics of a project were shown to affect the subject matter of emails. The purposes for sending emails however were found to be generally consistent between projects of different characteristics. This was shown by comparing the pilot and major case studies for which there was only 1% difference in the frequencies of *informing*, *management* and *problem solving* contributions.

A number of findings were identified in relation to distributed and collaborative working. Inter-Company emails were found to more often discuss the *product* than Intra-Company emails. For two people one sending inter-company and one sending intra-company the contrast was as much as 95% versus 55%. However the purposes for sending messages were found to be consistent between these two types. The location of the parties communicating was found to have an affect on the subjects which they chose to discuss via email, with distributed users sending approximately twice as many *product* related emails.

Personal preferences were found to affect how email is used, in terms of both the topics and purposes for which emails are sent. The job role of a user was also found to affect these, particularly demonstrated by site managers, where the means of communicating in such a role is constrained by location.

Throughout the lifecycle of a design projects, the use of email was found to change. The topics which emails relate to were shown to be reflective of the stages and activities taking place in the design process when they were created. Furthermore, the purposes of email sending were also found to vary depending on project successes and difficulties. Finally, it

was shown how purposes of emails could act as a signature of design activity and project progress, and hence be used as a project management tool.

## 9.0 Discussion

In this chapter the implications of the findings with respect to the broader organisational and management issues are discussed. This meets research **Objective 6**: *To highlight implications and opportunities to improve the way in which email is used to support engineering design projects.*

This Chapter draws together all available evidence to provide suggestions for improved email use in the engineering design context. Some of these are strongly founded based on the evidence presented in the previous chapters; others are supported but will require further research.

Some of the key suggestions and observations for improving email use discussed in the following sections are presented in Table 9-1. This provides a brief summary of the discussion and gives an indication of how strong the arguments for the suggestion are and to what extent further work is required.

Following Table 9-1, the first section of this chapter explores the implications for information management (IM), including aspects such as record keeping and information reuse. The second section then explores the results in the context of knowledge management (KM), discussing expertise location and knowledge sharing. Thirdly, issues relating to project management (PM) and how email impacts upon this are discussed. The fourth section explores the implications for collaborative working and an understanding of how successfully or otherwise email supports this. Finally implications for the practice of using email, including guidelines and user training are discussed.

**Table 9-1: A summary of key suggestions and observations to improve email use discussed in this Chapter.**

Area		Suggestions and Observations	*
Information Management	Record Keeping	Streamline records by removing low value emails from collections making searching easier.	B
		Information Managers would benefit from understanding what factors affect email use and hence archive content.	B
	Design Information	Email contains important design information managers should account for this.	B
		Email needs to be considered as part of holistic IM strategy and integrated with other sources.	A
		Companies should not assume all design information is captured in email. Use policy to mandate what email should and should not be used for concerning design.	B
	Classification and searching	Use simple 3-4 category categorisation at a high level to aid searching and retrieval by eliminating large portions of an email set.	A
		Automated search techniques would benefit from the context of email threads	B
		Use awareness of email use to assist searching. Factors include project phases, location of parties, role or parties and	A
Knowledge Management	Community Structure	Automated expertise location systems using email must account for limitations, factors affecting email use may distort perception of expertise.	A
	Knowledge Sharing	Unofficial groups and hierarchies developed through email networks will mean that knowledge is not shared as effectively as possible	C
		Email is not the best medium for developing knowledge and relationships between team members will determine how effectively it is used.	B
Project Management	Contractual Obligations	Employees should be trained as to the legal significance of email and company email policies should be clearer about what should and should not be committed to email.	B
	Effective Communication and Working	Email should be used more consistently and users could classify their own messages.	B
		Engineers should understand the benefits and disadvantages of using email for design work, and these should be investigated.	B
	Emails as a PM tool	Use changes in email traffic as signatures of success and difficulties to support the process of project management.	B
Collaborative Working	Inter-company Collaboration	Inter-company email standards should be set before projects.	C
	Effective working relationships	Develop and initiate relationships outside of the email environment to improve effective email use.	A
	Effective working in design	Emails are not the most effective medium for negotiation and reaching consensus and should not be used alone for designing.	A
Email Practice	Structured Composition	Structuring email composition would assist use, retrieval, effective working.	B
	Guidelines and Training	Guidelines and training should be generally improved to include structure, content, purposes and addressing of emails.	B
<p>* Strength of argument, where:</p> <p>A = strong argumentation or supporting argumentation exists.  B = further research would be beneficial.  C = further research required.</p>			

## **9.1. Information Management**

Design is regarded as a process of changing information states (Ullman, 1992; Hicks, et al., 2002). The effective management of information, design related or otherwise, is critical to the success of a project.

The findings identify a number of influences on the nature of information contained in emails. Awareness of these is important for information practitioners. In this section implications are discussed specifically in relation to; general record keeping, managing design information and systems for automated classification and search.

### **9.1.1. Record Keeping**

In both of the case studies it was found that the majority of the email record related to the project or company rather than the specifics of the designed artefact (6.4.1 and 7.4.1). Furthermore, these email collections were also found to include a large amount (25%) of material of an administrative nature (7.4.1). The importance of such material will depend partly on the characteristics of the project and also on the approaches and policies that govern which emails are retained. From undertaking the coding it was apparent that many of these types of email in the collections were likely to be of little further value.

It follows that more streamlined record collections could be created if material of low value could be identified at an early stage and filed separately. The streamlined collections would then be easier to search and manage, and more likely to provide useful information when required.

A broad range of factors have been shown to influence the subject matter and purposes of email and hence email records, such as the personal preferences of project participants (*finding 13*), their location (*finding 12*), project characteristics (*findings 8 and 9*) and a user's role (*finding 14*). In order to create effective IM systems, information managers would benefit

from appreciating and understanding these dynamics, particularly with respect to their own organisations.

The topics of emails, and hence type of information, captured at different stages of the project was found to vary (*finding 16*). For example, discussion of the product was predominant at the testing phase. An implication of this is that an approach to keep records in such a way that they may be searched relative to a project plan would aid retrieval and allow emails associated with particular phases to be retrieved.

### **9.1.2. Design Information**

Although much of the email record collections has been found to relate to the project and company, rather than the product, a number of emails still contain design information (*findings 5 and 6*). These emails may either share information to support design work taking place elsewhere (up to 23.1%) or act as the medium in which design activity explicitly takes place (2.5%). In both cases these emails form part of a design project record, and it is anticipated that a proportion of this information will not be recorded elsewhere (*finding 6*). It is therefore vital that project managers and information practitioners are aware that emails do contain design information, even if they do not believe this is good practice. They can then choose to either discourage this practice or account for it when collating and managing records.

The findings state that up to 23.1% of emails facilitate design work (*finding 5*), where the design rationale contained in these largely supports decisions taken elsewhere. Some of these emails will contain fragments of a discussion and can be difficult to reuse without supporting email threads or a broader context from outside of the email environment. These emails contain potentially vital design information which may be unintelligible without the link to other sources. The implication is that email needs to be considered as part of a holistic IM strategy and that the links between design related emails and other information sources must

be maintained. As part of this, managers should encourage engineers to be clear and consistent in how, when and where such information should be recorded.

At the outset of this research it had been asserted that design information created between project team members was being exchanged in emails and hence recorded, and that the major issue was how to effectively retrieve this information. The findings however, show that a range of factors affect email use in design projects (*finding 15*) and some members of a team may communicate very infrequently via email, yet may often work together on design elements in person or via the telephone (*findings 12 -14*). This has identified that there is now a major risk factor, where companies assume that information is being recorded via email but in fact these transactions are not taking place in this way. It is imperative that companies do not assume that all design information is being captured by email and must look widely at practices and policies that govern other communication to ensure that all necessary information is recorded. The concerns of the industrial partner raised in Chapter 1 reflect this.

It would be useful to have an understating and expectation of what information is only contained in email repositories. This might then be mandated against. An alternative would be to provide post processing, which may be automated to make emails easier to retrieve.

### **9.1.3. Classification, Search and Automation**

Having captured information and identified which aspects should be stored, the next step is to organise this information for potential searching and retrieval. Classification can provide a key part of this and can potentially be automated (Nenkova and Bagga, 2003; Koprinska, et al., 2007).

The literature discussed in Chapter 3 from the information retrieval (IR) community showed that basic automated classification techniques are most reliable when there are few (three to four) categories. However the range of material within each category is broader, providing less differentiation for searching than if more categories are used.

The approach adopted in this research utilised only a few top level categories as suggested by Lincoln and Guba (1985). The three categories were *product*, *project* and *company*. From the case study results these were shown to be appropriate for decomposing the content of the emails as well as being an effective way of distinguishing between emails (6.3.4). Disregarding the few emails which contained so little text or context as to make sense, every email examined could be classified using the scheme.

Based on these results categorising emails at a high level in three to four groups can offer a simple, reliable and effective way of aiding searching;. This approach could also be automated using language processing techniques such as rule based retrieval (Blumberg and Atre, 2003)<sup>5</sup>.

Such classifications would eliminate the emails such as company based which have been shown to not provide design information, enabling effective searching. For example if you were targeting product information for example for a vessel you may use key words such as 'navigation system' to yield the most appropriate email information

One disadvantage of this approach is that the findings and observations from the case studies show that emails often contain little text (6.1 and 7.1.2), the context for which is made clear by other emails, telephone calls or meetings. In these cases searching though the body and subject

---

<sup>5</sup> Personal communication with Dr Andy Dong, University of Sydney, in a meeting at the IMechE, London, February 2009.



of an email using keywords will miss potentially important messages, particularly if the search is speculative. The implications of this are that automated classification systems may need to search previous related emails in order to appropriately classify the next (as explored by Nenkova and Bagga, 2003). This could be achieved by utilising the thread features within email systems such as Google™ Mail (5.6.1).

Other findings from this research could also be used to assist speculative searching by providing increased awareness of how emails are used within a project. For example, knowing that more design information is exchanged in inter-company emails makes these a priority when looking for design information (*finding 10*). Knowing how different email topics and purposes change over time could also enable searchers to use timeframes intelligently; for example knowing that product costs tend to be discussed earlier in the project (*finding 18*). Finally associating users' roles with a relevance to certain types of information could also be used to facilitate searches (*finding 14*). The information gathered from the emails can then be used as part of an overall knowledge management approach within companies.

## **9.2. Knowledge Management**

Managing knowledge within an organisation is seen to be important in order to maximise effective working. It is recognised that email use is one of the ways in which participants generate and share knowledge with one another.

In this section, the implications of the findings in relation to knowledge management are discussed with respect to the identification of expertise in a company through email records and the effectiveness of emails as a mode of growing and sharing knowledge.

### **9.2.1. Community Structure Expertise Location**

As discussed in 3.2.2, research has already shown that it is feasible to locate expertise using networks created by email traffic and using keywords of content (Kim, 2002; Campbell, et al., 2003). The findings however have shown that a number of factors will impact on the way email is used by different people in different roles (*finding 14*). Hence, a map of expertise gained solely through using email traffic and content will not be a true reflection of the expertise within a company.

If knowledge managers have an understanding of which factors and to what extent they influence email use within a project or company (*findings 7 -15*), they should then be able to compensate when mapping expertise. If at the very least knowledge managers are aware of uncertainties and incompleteness in expertise maps, they can treat this resource with an appropriate degree of caution.

### **9.2.2. Knowledge Growth and Sharing: Is email an effective communication medium?**

The literature suggests that email is less effective for certain processes, such as reaching consensus over decisions and idea generation (Wilson, 2002). It also reports that the development of new understanding and creation of new ideas in emails is subservient to information transfer (O'Kane, et al., 2007) This concurs with the findings that few design ideas are generated and shared in emails (*finding 6*) whilst information transactions are common (*finding 1*). Email is not therefore the most appropriate medium for knowledge growth. This needs to be recognised by knowledge managers and engineers alike. When undertaking tasks users should then be better placed to consider whether email is the most appropriate or necessary means of communication. Furthermore the way in which emails are expressed will be influenced by the relationships between team members (*finding 7*). This means that the integrity and value of emails will vary. Hence although there is a knowledge sharing

benefit to email this will not be consistent across all projects or even a single project.

Email enables members of a project team to form their own structure of direct communication with colleagues from another company, and a manager may wish to review whether these are appropriate. Emails enable an unofficial chain of command and there is a risk of not utilising everyone in a team effectively if, for any one of the factors identified, key players are not involved in a network (Panteli, 2002). This will mean that knowledge is not shared or created as effectively as might be possible.

### **9.3. Project Management**

Emails are a vital project management (PM) tool, and used more so for this than for designing. Managers should recognise the significance that emails play in the control of a project and understand that key decisions or outcomes can be driven solely by email communication.

In this section the implications of the findings are discussed with regards to key aspects of PM; contractual obligations and effective communication and working. Finally the use of email as a tool for aiding PM is discussed.

#### **9.3.1. Contractual Obligations**

Inter-company emails within engineering projects potentially contain elements of contractual significance made unwittingly by employees. ***“Email now has an important central role in contractual situations and is required to evidence in any litigation where its content may be very important to establish design intention however unwittingly it was recorded in the email. ... The management implications of this***

***are not fully recognised [by organisations]***<sup>6</sup> The major case study (Chapter 7) showed a 32% of emails were sent between the industrial partner and external parties and hence this issue is clearly important.

A primary observation based on this is that employees should be trained to have greater appreciation of the legal significance of email; this might encourage a more considered approach to email content. Secondly, companies would benefit from having a clear policy on what should and should not be committed to email, for the purposes of project risk (AIIM International, 2009).

### ***9.3.2. Effective Communication and Working***

One of the key findings from the case studies is that personal preferences affect how email is used (*finding 13*). The resulting lack of consistency, between project team members, could result in incomplete evidence being collated via email, making work more difficult to conduct effectively. Managers should seek to set best practice standards encouraging consistent email use within projects. This might include the use of templates, structures and forms within emails and guidelines as to which communication method is most appropriate for different circumstances. The advantage of this for 'work related' emails is that the formal structuring would assist the automated analysis of the emails. This can be achieved by getting users to classify their own emails. An example of this is the 'contact us by email' facility provided by Sainsbury's<sup>7</sup>; where users are required to identify which of five topics their email relates to, this enables it to be routed correctly.

At present, the literature suggests that many companies do not have a clear email policy, particularly regarding appropriate composition and

---

<sup>6</sup> Personal communication with Dr L D Burrow, Converteam UK Ltd, 24<sup>th</sup> July 2009.

<sup>7</sup> [http://www.sainsburys.co.uk/sol/contact\\_us/contact\\_us.jsp](http://www.sainsburys.co.uk/sol/contact_us/contact_us.jsp) (accessed 1st June 2010).

content of emails (AIIM International, 2003; 2009). The interviewees from the case studies had differing views on what emails should be used for and when. The results tended to support this, showing differences in purposes for email use between the many participants of the project (*findings 13 and 14*). It is therefore recommended that along with encouraging project managers to standardise email through improved practices in their project, companies would benefit from enacting a comprehensive email policy. This could include provided a structured approach on how work emails are classified.

A further benefit of encouraging consistent use of email could be a reduction in email overload (Bellotti, et al., 2005). From both the interviews and observations it appears that many emails are unnecessary. A more structured approach to email use may help to reduce these in number.

It was found that design work can take place within emails (*finding 6*). The literature (e.g. Wilson, 2002) suggests that emails are not an effective forum for certain behaviours used in design, such as reaching consensus or decision making. Whilst it may not be practicable to avoid using emails for designing, it would be beneficial for engineers to understand their strengths and weaknesses for this, enabling them to use emails most effectively.

In Chapter 2 design was discussed as a process of transferring information from one state to another. The findings show that emails are particularly important for this information transfer at the inter-company boundary (*finding 1*). In order to improve the effective working with emails for design, it is vital that the terms for inter-company email exchange are clear. Examples of terms could be: making the subject line of an email include the relevant sub-system reference number, using separate emails for discussing separate problems, marking up the subject line of administrative emails with an 'A', not 'cc-ing' people into emails without explaining why in the body of the email. It may be desirable to consider

such standards for email use between parties at the outset of a project, in order that information can be shared most effectively. In practice this would mean all parties coming to an agreement on practices, terminology and referencing that they would all follow.

### **9.3.3. *Emails as a Project Management Tool***

The purposes for sending emails were found to change over the duration of the project. The trends for these changes were shown to represent signatures for different circumstances encountered (*Finding 17*). For example, steep increases in problem solving activity, along with increased information transactions and management highlighted difficulties encountered in the engineering work. An increase in information and management transactions with a decrease in problem solving characterised a project in a difficult phase.

Managers could, by understanding these, have an expectation of how email should be used and hence determine whether a project is proceeding successfully or not and what sort of intervention might be required. For example where managers find that problem solving is increasing towards the end of a phase, this should be recognised as a cause for concern and appropriate action can then be taken to avoid missing a deadline.

Although the suggested project related reasoning for the email traffic signatures seen in the major case study may not accurately identify all project situations (8.5, Table 8-3), their alignment with actual events in this case study suggests the value of email information in a group context. This means that a group's email content can be used to provide valuable PM insight, without the requirement to identify individual's private information. If project managers could recognise these signatures and correctly interpret them for their circumstances, it may be possible for them to take intervention action sooner rather than later. The idea of using group information in this way is often referred to as 'The Wisdom of

Crowds' (Surowiecki, 2004). An example of this was the co-ordination of relief efforts and identifying where resources were most required following the Hati earthquake, using mobile phone texts from rescue and relief workers (Financial Times, 2010).

Through exploring the purposes of a groups' email sending, the presence of a signature can suggest there may be a problem forming. Based on the baseline email traffic, deviations from the traffic, and in particular the rate of increase or decrease from the baseline, this could help to identify whether a problem is emerging. Second, these traffic patterns (Table 8-3), could help to identify opportunities for reflection. For example, during the testing phase, when difficulties did occur, emails relating to problem solving increase dramatically. This suggests that email records from this period may contain valuable lessons to be learned, pointing to both issues and resolutions.

## **9.4. Collaboration**

As a convenient means of communication between two or more parties, emails are a useful tool for collaborative working. It is therefore important to understand to what extent emails are an effective means for collaborative working.

In this section the implications of the studies are discussed with regards to effective working and then collaborative design. First, because of its significance for today's globally distributed teams, inter-company collaboration is discussed.

### **9.4.1. *Inter-company collaboration***

Each company may have its own standards and policies towards email use or a lack thereof (AIIM International, 2009). This may include conventions for subject line titling and inclusion of reference numbers. Such inconsistencies can make classification and archiving more difficult

for both parties. The findings show that the purposes for sending emails are very much consistent between inter and intra-company messages (*finding 11*). It is suggested that prior to the start of a project, conventions should be mutually agreed by the major parties involved to incorporate their needs. This could also go as far as including standard layouts of messages and topics and purposes for which email is or is not suitable. This practice could also encourage group mail boxes to be used for queries not specific to an individual engineer which can then be more efficiently dealt with and responded to (The Radici Group Inc., 2008).

The findings also highlighted that there is an increased discussion of the product and that much design work is acted and facilitated via email at the inter-company interface (*finding 10*). This shows that email is a critical tool for designing as well as for PM. It is therefore vital that these emails are clear and effective, otherwise the success of the design may be compromised.

Email is a popular medium for communicating inter-company because of its low cost, convenience of use, and asynchronous and time zone benefits in comparison to mediums such as phone and fax (Wilson, 2002). This means that companies will struggle to restrict email use, for the purpose of controlling design information, without compromising the key benefits it affords.

In practice, organisations should accept that emails are now a part of inter-company communications and treat them as a key resource. Through ensuring that workers are appropriately trained, email records should then be useful and not compromising contractually.

#### **9.4.2. Effective Working Relationships**

A key aspect of collaborative working is the forming of relationships between new members of a project team, who may not previously have met or worked together and who may be geographically dispersed



(discussed in 2.4). It is important for these people to form effective working relationships in order for project success. It is suggested that email is not an ideal medium for introductions and initiating relationships (Wilson, 2002).

In the case studies it was found that the way in which emails were expressed was influenced by the relationships between team members, with closer groups expressing more personal, particularly positive emotions (*finding 7*). This supports the literature which suggests that communication processes are socially orientated (Eckert and Stacey, 2001). This suggests that users who already know one another are happy to discuss matters openly and honestly in emails and should understand each other and work well as a team.

The implication of this is that, where possible, relationships outside of the electronic working environment should be initiated and reinforced prior to a project. This should help to improve the closeness of the parties involved. This in turn should improve the effectiveness of their team working and email communication.

A further issue for effective working may be caused by the differing approaches to email users (*finding 13*). The findings show, for example, that people in similar roles may use emails in different ways according to their personal preferences. This invariably means that some people will be more involved in email communication than others. This may have the effect of causing sub groups to form within a team, which might not be apparent to the manager. The risk therefore, is that the team will not function as an effective unit.

#### **9.4.3. Effective Working in Design**

Three high level design communication scenarios were identified by Eckert et al., (Eckert, et al., 2005) namely handover, joint design and interface negotiation. As discussed in *finding 1*, email was found to be

widely used for informing, supporting the handover scenario. By definition joint design (co-located and synchronous) is not supported by email.

Emails were seen to support interface negotiation (*finding 5*). Indicators of this were the use of terms such as; requests for information, solution development, clarifications and exploratory discussion. The output of an interface negotiation however is a decision or consensus, which is rarely found in emails. This identifies that emails must be considered alongside other communication mediums in order to obtain a complete picture of design communication.

Existing literature also suggests that email is not a good medium for negotiation nor reaching consensus (Wilson, 2002) both of which are required when making design decisions as part of a team and concluding interface negotiations.

The case studies show that, in practice, few decisions are clearly made and expressed in emails (*finding 4*). However emails were often used for exploratory discussion and solution development, both of which are evidence of problem solving (*finding 3*). These creative elements are also foundations for later decisions and support the view that email is used for decision making purposes (Lusk, 2006).

The findings from the study imply that lots of rationale for decision making is present in emails but not explicitly linked to the decisions outcome. Where a number of emails, possibly between different parties, relate to one decision this may not be clear. Not only is there a record keeping implication for this, but team members will not be able to work as effectively if they are not party to a comprehensive set of information (Eckert, et al., 2001). It would therefore be beneficial to associate emails with the tasks they relate to, making all the reasoning and evidence more explicit. This should include ways of making this information accessible to all of the relevant players within a project. This could include using wikis for discussions so that other project members could contribute or observe

discussions, this would also have the advantage of maintaining threading. Another option is to support the use of a shared email repository, which project members should be encouraged to file messages in as well as review existing content.

## **9.5. Email Practice**

Emails are used idiosyncratically making their management a difficult task (3.1 and finding 13). It is suggested that improved practices may help to control and standardise email use, aiding information management (IM) aspects.

One possible route to improving email practice is to structure the composition of emails to a greater or lesser extent; i.e. fixed templates for basic emails through to criteria for the key pieces of information an email should contain. Implications of the findings to this route are discussed. Following this, implications for the guidelines and training given to engineering email users are presented.

### **9.5.1. Structured composition**

One possible route to improving the use and reuse of emails is to compose them around a predetermined structure. Any kind of structure needs to encompass the variety of purposes and topics for which emails are sent, as demonstrated in the case studies, with out being excessively restrictive.

It is feasible that an equal number of important emails will have originated from outside of the organisation as from within. This means that for a structure to be truly beneficial to a project all parties should adopt it. Research into structuring email composition should therefore investigate how this might be achieved.

It was found that around a quarter of emails fulfilled administrative purposes (7.4.1). Within this type of email, common purposes were often observed, such as booking a hotel room or asking for an invoice. It is suggested that standardising such requests might be part of an overall approach for improving email use via structuring as discussed in section 9.3.2 above and discussed further in 10.7.1.

It was generally observed that most requests and objectives of emails were often entwined in discussion and not clearly flagged. As such, these messages may not be as effectively communicated as if they were to be broken down to fit around an enforced structure. A trade off or compromise may therefore be necessary.

#### **9.5.2. Guidelines and Training**

Many organisations now have some form of guidelines in place regarding email use (AIIM International, 2006) which will usually be tailored to meet their needs. The findings highlight some aspects which these would be beneficial to consider.

The findings show that it is common place for emails to have more than one purpose or subject of discussion (7.4.1). Instead of sending one broad ranging email to a person, users could send separate emails to the same person addressing each area. Users should be directed to consider when this would be of benefit for the storing, organising and re-using of emails as well as to communicate more effectively (see 2.6.2).

The findings of the case studies (*finding 1*) and wider literature (2.6) show that information transfer is a critical part of the design process and one of the dominant roles of email. The implication of this is that education of engineers should cover effective information sharing including email use, because of the significance of this to the design process and PM.

The breadth in purposes and topics of emails from the case studies illustrated the numerous ways in which emails can impact on projects and the variety of their relevance and significance (6.4.1 and 7.4.1). It is suggested that users should have a clear understanding of the potential uses and implications of their emails. This should enable them to appreciate the benefits of following company guidelines and protocols. If IM approaches beyond 'keeping everything in sight' are to be successfully adopted this will require users to better understand the role that email plays.

Email use varied greatly between participants in the case studies and it was shown that a number of factors including personal preferences impacted upon this (*findings 12 - 14*). Encouraging email use by those who would otherwise not use email and encouraging all project members to use it in a similar way would improve the overall consistency of use within a project. Consistent and complete email records should help facilitate IM and KM.

As a final point it is generally suggested that matters of email etiquette, such as appropriate use of the *to* and *cc* fields, could be better practiced by users. It was observed in the case studies, interviews and wider discussions, that such basics are not followed.

## **9.6. Summary**

This chapter has identified implications of email use in engineering design projects by considering findings from the case studies and existing literature. The areas of information management, knowledge management, project management, collaboration and email practice were discussed in turn. In doing so **Objective 6** has been met.: *To highlight implications and opportunities to improve the way in which email is used to support engineering design projects.*

This chapter began by discussing the implications of the case study findings to information management. It was highlighted that Information Managers should understand the dynamics within a project which influence email use, such as distributed working and participants. They should also be aware that emails do contain potentially important design information and because these often support decisions made elsewhere, emails should be integrated with wider records. It was further suggested that an initial basic manual categorisation of emails into three to four groups could assist in sifting through emails, when used in parallel with conventional search techniques and filters.

Knowledge Management was then discussed. In support of the existing literature it was identified that expertise might be located and mapped via email content and traffic. Because of differences between users' email behaviour, such maps will not present a complete picture of project or company knowledge. With regards to knowledge growth and sharing; users must consider whether email is the most appropriate or necessary mode of communication for the task in hand.

Implications for project management were discussed. Amongst these it was identified that emails can form unwitting contractual obligations, which employees should consider agreeing when composing inter-company emails. It was also suggested that standards for email use across the parties in a project would improve consistency and make working and communicating more effective. It was also shown that changes in email use could be a potential tool for Project Managers to identify progress and possible issues in a project.

Collaborative aspects of working via email were then discussed. With regards to inter-company collaboration it was suggested that distributed teams should meet in person, so that they can form stronger bonds with one another. This should help them to communicate more easily and openly via email. With regards to collaborative working generally, it was identified that emails need to be better integrated with other information

sources. This is primarily because emails contain rationale which is not explicitly linked to decisions that take place elsewhere.

The final section discussed implications for directing best practice in email use. It was suggested that a few basic types of message which are often repeated would be easier to compose and read if templates were used. Applying and following protocols such as when to use the *to* and *cc* fields make emails easier to read, record and organise. It was suggested that engineering education should have more emphasis on information management, because it is so critical to project success. Having this understanding might encourage engineers to invest more time in creating better emails. Finally, managers should encourage email users to be more consistent with one another in how they use email.

## 10.0 Conclusions

The overall aim of this research is to investigate the role and use of email in engineering design projects through an analysis of email content. For this research the role and use of email comprises the purpose for sending emails, the subject matter they relate to and the activities they are used to support.

To meet this aim three research questions were posed, namely:

**Research Question 1:** What is the nature of the information engineering teams communicate using email?

**Research Question 2:** How does email support project management and design activity?

**Research Question 3:** How can an understanding of the content be used to support engineering design projects?

Prior to answering these, the following section discusses the overall contribution of this research. A summary of the key findings is then presented. The research questions are then discussed in turn based on the findings of the research undertaken.

A further section then provides a summary of the work undertaken and the objectives it has met. The concluding section suggests direction for future research, including developmental points arising from the methodologies adopted in this project and a reflection on the coding scheme.



## **10.1. Contribution to Knowledge**

This thesis makes two significant contributions to engineering science. The first concerns the creation of an approach for the analysis of the content and use of email in engineering projects and teams. The second contribution concerns the analysis of the email content of a large engineering project and its relationship to project performance. Each of these contributions is summarised and aspects are discussed in more detail in subsequent sections.

### **10.1.1.        *Content Based Approach***

A new approach has been created to enable the textual analysis of the content of engineering email. The approach considers *what* topics emails include, *why* they are sent and *how* their content is expressed. This has been developed from existing literature and iterative application to emails from a variety of engineering projects. Whilst in this work the approach has been developed for email analysis, other researchers have applied derived methods to other engineering information sources such as log books (McAlpine, 2010). The approach may therefore offer a more generic process for characterising the content of engineering information.

### **10.1.2.        *Analysis of Emails and Content***

The analysis of the content of email in engineering has been achieved by coding a sample of emails from a large systems engineering project (comprising 16 000 emails). This was achieved by using the aforementioned approach to examine the subject matter of emails (*product, project, and company*) and the purpose for sending emails (*information, management and problem solving*). Based on this analysis a generalised perspective of email use in engineering design projects has been established. Further by considering the time phase analysis of content, the different relationships between changing content and project performance have been identified. In particular this includes the relative volumes concerning information, management and problem solving.

## 10.2. Key Findings

Chapter 8 analysed the results from the coding of email corpuses from the two case studies (Chapters 6 and 7); eighteen key findings were identified. These are shown in Table 10-1, and referred to in the following sections.

**Table 10-1: A summary of the key findings from the two case studies.**

1	The predominant purpose of emails is to transact information
2	Emails are used for directing activity and management
3	Emails are used for exploratory discussion and solution development, both of which are evidence of problem solving.
4	Few decisions are clearly made and expressed in emails.
5	Design work is facilitated via email
6	Design work takes place within emails
7	The way in which emails are expressed will be influenced by the relationships between team members
8	The characteristics of a project will affect the subject matter of emails
9	The characteristics of a project have a limited affect on the purposes for sending emails
10	Inter-Company emails often discuss the product more than Intra-Company emails.
11	The purposes for sending emails are consistent between inter-company and intra-company messages.
12	Location of parties affects the subject matter of emails
13	Personal preferences affect how email is used
14	A user's job role has an affect on the topics and purposes of their email
15	Combinations of factors affect the topics and purposes of email
16	The topics of emails reflect the stages and activities of a design project
17	Email purposes can be signatures of design activity and project progress
18	Email use is dynamic

### **10.3. Research Question 1: What is the nature of the information engineering teams communicate using email?**

An approach was developed (Chapters 4 and 5) which identified the information within emails as relating to a *product*, *project* or *company* and facets thereof. An email sending purpose was also identified under the groupings of *problem solving contributions*, *information transactions* and *management transactions*. With regards to the research question, the nature of an email is considered to consist of a subject matter and a purpose. The nature of emails was also found to change over project duration. These three concepts are now discussed.

#### **10.3.1. Subject Matter**

The subject matter of engineering project emails can be identified by any combination of the three key topics they relate to, the product, the project and the company. In case studies (Chapters 6 and 7) from two systems engineering projects of differing characteristics it was shown that these topics varied (*Finding 8*). Product related emails were sent relatively consistently, being referred to in 35% of all emails in each project. However the proportion of emails relating to the company varied between less than 2% and over 50% between the different projects. It is concluded that the characteristics of a project (such as domain and design type) will have a significant impact on the subject matter of the information contained within emails.

Variation in the subject matter of emails also depended on whether emails were sent inter-company or intra-company. With inter-company emails more often found to relate to aspects of the product design, 95% to 55% in one example (*Finding 10*). It is reasoned that this is because it is easier to discuss design elements *face-to-face* where possible, which is

more often the case in intra-company communication where participants are co-located.

### **10.3.2. Purpose**

The purpose of emails in these case studies, was most often found to be to transact information, approximately 40% of emails were sent solely for this purpose (*Finding 1*). A thorough review of email literature (Chapter 3) indentified that information transfer was likely to be the most common purpose of email sending, concurring with this finding. Emails relating to management purposes occurred less often with between 16% and 19% of emails sent solely for this purpose. The least common purpose for emails was to achieve the problem solving activities associated with design, 6% of emails were sent for this sole purpose. It was found that emails are often used for multiple purposes, with between 41% and 49% of emails fulfilling two to three purposes. In contrast to the topic or subject matter of emails, the purposes for email sending were found to be consistent between the projects, which varied in terms of the domain, design type, duration, budget, geographical distribution and range of involved parties (*Finding 9*).

The nature of information communicated via email was shown to be different for parties with different roles in the design project (*Finding 14*). Directors for example were found to send 40% of management related emails compared with 10% for site managers who instead communicate more for problem solving. The personal preferences for modes of communication and styles of working were identified as a significant further factor in the differences between how individuals contributed to the teams email communication (*Finding 13*). In interview some participants stated they preferred communicating *face-to-face* whilst others preferred the convenience of email.

### **10.3.3. Temporal Element**

It was found that the nature of information shared via email varied over the life of projects. In particular the topics of emails were seen to reflect the corresponding stages of activity which were taking place (*Finding 16*). For example product performance was discussed five times as often during the testing stage compared to the assembly stage. The extent to which emails were used to manage tasks and facilitate problem solving also corresponded with the needs and pressures of projects at key times.

## **10.4. Research Question 2: How does email support project management and design activity?**

The review of engineering design literature (Chapter 2) identified activities and requirements in the design process and project management. This was used along with the review of existing email literature to analyse the case studies to identify how effectively email supports engineering activities (Chapter 8). Engineering emails primarily support two activities, designing and project management (*Finding 2*).

### **10.4.1. Project Management Activity**

It was found that emails were commonly used for project management. This was shown by the 45% of emails which fulfil a management transaction and in particular the 34% of emails which fulfil a management transaction and contain the project in the subject matter. It was also found that emails are used not only by managers to direct activity, but by engineers to request actions of their managers and peers.

The high volume of project management emails is evidence in itself that email is perceived to be a convenient way of conducting such activity, for the sender at least; recipients however can feel unnecessarily over pressured to respond (Renaud, et al., 2006). A major pitfall of using email as a project management tool is the ease with which high work loads can

be generated by overloading of inboxes (Whittaker and Sidner, 1996; Whittaker, et al., 2006). Evidence from the case studies supports this concern; with high volumes of email sent and received particularly by the project director. Project participants also expressed concern about the amount of email they had to deal with.

Existing literature also highlights the potential for employees to make unwitting contractual commitments by making suggestions or agreements in emails, which can then be used as written evidence (Genborg, 2005). This issue was further raised by the industrial partner company to this project. It is argued that company policies and practices regarding what emails may and may not contain would reduce this issue. Currently this policy is usually lacking (Allim International, 2003) and as such the case studies highlighted inconsistencies in the ways in which email was used by different members of the project team.

#### **10.4.2.        *Design Activity***

Emails supporting design activity were identified by having a product related subject matter, these were less frequent than project emails in the case studies. When examining the way in which email is used to support design activity a distinction was been made between facilitating work via email and conducting work in emails. Product emails with an information or management purpose accounted for a fifth of email sending and facilitated design work (*Finding 5*). A mere fortieth of emails contained problem solving purposes related to the product which can be considered to represent 'designing' in email (*Finding 6*). Whilst this does not sound significant it corresponds to more than 400 emails in a project the size of the major case study, which equates for around 8 000 lines of information.

It was also found that emails were used for general problem solving activity, not specifically related to the product or design (*Finding 3*). Although few decisions were explicitly made in email content (*Finding 4*),

elements of exploratory discussion and solution development were often present, evidence that email is used to support decision making processes and activity.

The approach also explored how email content was expressed, along with existing literature exploring differences between email, *face-to-face* and other communication media. It is argued that any disadvantages of using email are likely to be most pronounced where design work takes place in email. This is because email is less suited to negotiating and reaching consensus (Wilson, 2002) which are key in design work; indeed the case study found few decisions were made in emails.

On a day-to-day basis email can be effective at both facilitating and conducting design work. However the elements of design rationale, captured in emails, are often abstracted from their corresponding decisions, meaning that information management further down the line is adversely effected.

The practicality of communication that emails affords, particularly in projects where parties are geographically separated, may outweigh any disadvantages, so long as limitations are understood and other communication means are not rejected. Limitations revolve around social aspects of communication and the aforementioned decision making. Ideally parties should meet in person to form a cohesive team or working relationship. The potential for email overload can also be a disadvantage.

#### **10.4.3.        *Signatures of Project Progress and Design Activity***

It was identified that the way in which email supports project management and design activity changes over the duration of a project. Email enables the differing discussions required at each stage, for example about product costs early on and project costs later. In particular emails fulfil different purposes needed to drive projects forward, through improving

information transfer, enabling managing activities or facilitating problem solving.

Ultimately observations regarding the changes in the purposes for sending emails were identified as signatures of project progress and design activity (*Finding 17*). Discussion of how these signatures could improve the support of engineering projects is made in the next section.

### **10.5. Research Question 3: How can an understanding of the content be used to support engineering design projects?**

The implications of the findings (Chapter 9) highlight a number of key opportunities to improve information management and the efficiency of email use. In answer to the research question, the following recommendations are made.

#### **10.5.1. Information Management**

A structured approach to disseminate between relevant, irrelevant, and significant and insignificant material at an early stage in the record keeping process, would enable more streamlined records to be kept and searched. In tandem with this, categorising emails at a high level in two to three groups could be a simple, reliable and effective way of aiding searching. Examples of such categories are the product, project and company terms identified in this research.

While the extent of design information was not as much as expected, the importance of what is captured cannot be underestimated, hence the need to keep design records. However, email is informal and often lacks sufficient context to stand alone. Hence there is a need for it to be managed and integrated with both formal and informal records.



### **10.5.2.      *Knowledge Management***

There is potential for records of email traffic to be used to produce expertise maps within a project or organisation. These should be treated with appropriate caution and an understanding of which users may or may not be fairly reflected by them. This is because key players are not necessarily always active contributors in email networks.

Further to this email is used to maintain personal and extra-organisational relationships which could be important for knowledge management.

### **10.5.3.      *Project Management***

Project management would be more effective if those composing emails had a better appreciation of their significance and there was less variation between users' behaviour. Employee training could highlight the risk of making unwitting contractual commitments via email. Engineers should be supported in identifying appropriate scenarios for email use and working towards standardised formats for structure and content, thus reducing the variation and inconsistencies between the habits of individual users. This could be achieved by working towards a policy or standard for either their company or ideally, at an intra-company level for each project. Further training could highlight the potential strengths and weaknesses of email, particularly for designing and problem solving.

The research identified that 'signatures' of email activity could help to detect project progress and identify problems arising. This would merit further exploration and could prove a useful tool to project managers for determining whether their project is on track. The signatures were based on the relative volume of information, management and problem solving purposes in emails. Using data about the project allowed the relative changes to be associated with the project progress.

#### **10.5.4. Collaboration**

Parties collaborating via email would have more effective working relationships if they met *face-to-face* for an introduction at the very least. The strength of their team building skills outside of the email environment should transfer to inside it.

Where possible group mail boxes could be used, these should improve efficiency and prevent individual members within a team becoming overloaded. It also follows that the standardising of email practices, suggested to improve project management should also benefit collaborative efforts.

#### **10.5.5. Email practice**

Improved email practice has the potential to make significant improvements to project, information management and knowledge management and effective collaboration, as already discussed. With consideration as to the nature of their work and information requirements; it is suggested that companies form guidelines, provide training and encourage best practice by their email users, concerning when to use email, content and structure, and relationships to other information sources.

### **10.6. Project Objectives**

This research set out to investigate the role and use of email in engineering projects where very little research has previously been conducted. This was perceived as necessary due to the increasing prevalence of email in engineering projects, where effective Information Management and Project Management on a large scale are increasingly critical.

To achieve this three research questions have been discussed. In addressing these research questions, six objectives were developed:

- 1) To examine the existing research and determine what is known about the role that email plays both in engineering design and other working environments.**

This objective was primarily addressed in Chapters 2 and 3. Literature reviewed in Chapter 2 provided an overview of engineering design and systems engineering projects. It also explored how communication and information sharing supports these and consequent implications for email use. In Chapter 3, literature relating to email use from the disciplines of Human Computer Interaction, Management, Social Science, Language Processing and Information Retrieval was reviewed. This identified what is already known about the working processes of dealing with email and the secondary use of email data to support information and knowledge management.

- 2) To examine the tools and techniques proposed to analyse email use, including archiving and retrieval as well as general work processes.**

Chapter 3 also discussed the existing tools. It identified that most previous studies of email have used statistical data such as the frequency of messages and networks between communicating parties. Some tools were also found to have identified frequent key words within email messages to help classify them for archiving purposes.

- 3) To develop and validate an approach for analysing the role and use of email within the context of engineering design projects and the overall design process.**

Chapter 4 examined relevant methods from the engineering design domain which could be used to analyse email use in engineering projects.

Based on this, and the email literature previously discussed an approach was proposed to analyse email use in engineering projects. This centred on coding the content of emails from an engineering project identifying what topics they contained, why they were sent and how their content was expressed. This was to then be analysed for frequency of occurrence, variety of use between different parties in the project and changes in the email use over time. This was to be reviewed in relation to project documentation and the existing literature.

A key requirement for this was the development of a coding scheme, reported in Chapter 5. Principles of coding scheme development were identified from literature including, scope, reliability and validity. The coding scheme was developed by a group of researchers following an iterative process. The final scheme was validated by an application to 200 emails by three researchers.

**4) To apply the proposed approach to recently completed industrial projects and to explore the role that email plays over the project life.**

Chapter 6 reported on the application of the approach to a relatively small intra-company software design project. This pilot case study was used to further validate the approach, and provided initial findings. Chapter 7 reported on the application of the approach to a multimillion pound large scale systems engineering project involving numerous disciplines, companies and geographical locations, producing 16 000 emails over four years. Initial statistical findings such as the frequency of occurrence of each coding term were presented. These were also shown in relation to different senders and changing with time.

**5) To elicit general observations regarding the role and use of emails in engineering projects.**

Chapter 8 analysed the findings of the two case studies further identifying and explaining trends. The major and pilot case studies were compared and contrasted to identify general features of emails in engineering projects and factors which affect email use. The email literature from Chapter 3 was used to support this. Project documentation was also examined in relation to the findings, for example identifying how email use changes during different phases of a project. Overall eighteen findings were made.

**6) To highlight implications and opportunities to improve the way in which email is used to support engineering design projects.**

Chapter 9 brought together the existing research and current position of email use in engineering project with the analysis provided in Chapter 8. Implications of the findings made in Chapters 6 and 7 were discussed for Information Management, Knowledge Management, Project Management, collaboration and email practice.

Having achieved the objectives it is intended that this project should provide the understanding and direction for future work.

## **10.7. Further Work and Research Opportunities**

Future work is considered in three strands. Firstly, possible developments to the coding scheme and reflections are presented. Secondly, the ways in which this study could be further developed and the approach improved are discussed. Finally the potential research opportunities, identified as an outcome of this research are identified.

### **10.7.1.        *Development of the Coding Scheme***

An evaluation of the scheme was presented in 5.4, informing minor revisions. The final scheme was then applied in the case studies discussed in Chapters 6 and 7. Validation of the scheme was presented in 6.3. It may be that for further case studies it is desirable to expand or vary elements of the coding scheme to encompass new concepts or activities. In this case it is suggested that the method followed for deriving the existing coding scheme (Chapter 5) be repeated. Such opportunities for further coding scheme developments in future work are now highlighted.

#### **Topic Coding**

The topic aspect of the coding scheme was found to be fit for purpose (6.3). It is suggested that to develop the coding scheme further, the most frequently used terms could be sub divided to provide more insight; features and administration were widely used, 16% and 25% respectively.

#### **Purpose Coding**

Seventy three percent of emails in the major case study were sent for the purpose of an information transaction, most of these (60%) were sent for informing. It is suggested that there is scope here for development of the scheme. It would be beneficial to break down the types of informing further.

The literature and results indicate that emails are used for designing and to facilitate designing and consequently contain important design information and reasoning (*findings 5 and 6*). Whilst problem solving contributions were widely identified, the coding scheme did not link these to decisions or identify the development of arguments, (9.1.2). Less decision making was found than had generally been expected (*finding 4*).

It follows that the area of problem solving should be further explored. Of particular interest is the build up of designing reasoning. This could be

achieved by exploring emails at a lower level and in particular how problem solving discussions build through threads. An argumentation based coding approach may help to achieve this identifying just four concepts namely; issues, possible solutions, arguments and decisions. The coding could also be conducted entirely at the clause level to ensure all details are captured and that generalisations as to the email purpose to not obscure small but key elements such as decisions.

### **New Coding Terms**

During the coding process a few possible further terms were identified, presented in Table 10-2. These are not included in the results because they were not applied to the whole corpus or robustly defined. The frequency with which they continued to be identified is listed. Some of the concepts such as 'conflict within the team' will have been coded in any case by another label such as the Bales category in this case, shows tension. These terms are listed for consideration in future refinements or scheme developments. Terms such as 'contact details for somebody' do not readily fit into the topic and purpose approach of coding, but may provide a useful standalone statistic.

**Table 10-2 Possible further terms identified during coding**

Emails concerning...	Frequency
Something of urgency	4
Details of a warranty issue	3
Contact details for somebody	16
Conflict within the team	8
Customer or public relations	12
Identifying mistakes that have been made	13
Safety	7
Failures in service	4

## **Attachments and Threading**

The coding scheme and overall approach did not explicitly consider the use of attachments in emails (5.6). It follows that future work should explore the role and use of attachments in emails, including an analysis of their content. The coding approach also did not examine threading, and particularly the development of conversations at a micro level in threads. As previously discussed, this may require the coding scheme to be developed with an argumentation based approach.

### **10.7.2.        *Development of the Study***

The two case studies considered for this thesis contrasted in terms of project size and the geographic spread of parties. This has helped to provide a sufficient range of data in order to begin drawing general conclusions about email use in engineering projects. However it would be useful to spread further by applying the coding scheme to email corpuses for the different types of design activity, original, adaptive variant and for projects in different engineering domains (e.g. mechanical, civil, electrical).

Comparing the case studies it was noted that the proportion of Company related emails was far higher in the intra-company case study (*finding 8*). It was contended that the inter-company versus intra-company nature, of the pilot case study compared with the major case study, explained the wide discrepancy in the use of the *company* topic. Further work should explore in more detail the differences between inter-company and intra-company projects and explore this theory.

This research identified that there are significant changes in email use in engineering projects over time and with respect to different project participants. To provide sufficient data to analyse these aspects more thoroughly, future studies would benefit from sampling emails more frequently across a corpus



The partner company for this project reported benefits from the insight provided by having someone unfamiliar with their company practices looking in and providing comment and (macro level) analysis on their email use<sup>8</sup>. In return for such benefit, it is hoped that in the future other companies may be able to participate in studies, in order to produce a broader ranging picture of email use.

### **10.7.3.      *Research Opportunities***

The changing use of email over the life of a project has been discussed as a signature for project progress and design activity. This concept would merit further exploration, as using email as a project progress prediction tool could be of great value to project managers.

It has been identified that emails contain potentially vital design information, but that this often relates to other information sources. There is therefore an opportunity to improve the integration of emails with other such records.

The opportunity to use emails as a way of locating expertise, discussed in Chapter 3, remains. Caution should be paid however to the influence of a user's role and personal preferences over their email use. Further investigation of these factors in this context would be merited.

A number of automated techniques could be applied to automatically categorise emails into a few key groups. Either a rule based or retrieval based system (discussed in Chapter 3) could be used. Such techniques to classify emails in accordance with the coding scheme (or a development of it) would enable results in further case studies of large corpuses to be generated automatically. The principal difficulty is that

---

<sup>8</sup> Project review meeting with University of Bath and managers of Converteam UK, 26 February 2009, Rugby UK.

automation will need to identify where an email fits into more than one category. If developed, such a system could also be used to file emails in engineering projects, improving information management.

This thesis has reiterated the benefits of encouraging more consistent email use amongst engineers to help with information management and effective project management. There is a research opportunity to explore and develop ways of standardising formats of emails and providing guidelines for their use. This will need to assess practicality, as well as longer term benefits for information management and knowledge management.

Finally there is an opportunity to develop appropriate training, specifically for engineers, in order to make them aware of how important their emails are in the context of a project or even company and therefore how to best use them.

## **References**

The Oxford Dictionary of English, 2005. 2nd Edition Revised, Oxford: Oxford University Press.

AERY, M. and CHAKRAVARTHY, S., 2005. EMailSift: Email classification based on structure and content. *IEEE International Conference on Data Mining*, 27th - 20th September 2005 Houston, TX, United States. Piscataway, New Jersey, United States: Institute of Electrical and Electronics Engineers Inc., , 18-25.

AHMED, S. and WALLACE, K., 2003. Indexing Design Knowledge Based Upon Descriptions of Design Process. *International Conference on Engineering Design ICED 03*, August 19-21, 2003 Stockholm.

AIIM INTERNATIONAL, 2003. Email Policies and Practices: An Industry Study Conducted by AIIM International and Kahn Consulting, Inc., *Industry Watch*.

AIIM INTERNATIONAL, 2006. Email Management: An oxymoron? An Industry Study Conducted by AIIM International and Tower Software, *Industry Watch*.

AIIM INTERNATIONAL, 2009. Email Management: The good, the bad and the ugly, *An Industry Study Conducted by AIIM International, Industry Watch*.

AIRBUS, 2009. *Airbus World Wide* [online] Available from: <http://www.airbus.com/en/worldwide/>, [accessed 16/11/09].

AUGHENBAUGH, J. M. and PAREDIS, C. J. J., 2004. The Role and Limitations of Modelling and Simulation in Systems Design. *ASME International Mechanical Engineering Congress and RD&D Expo*, November 13-19 Anaheim, California USA.

AURISICCHIO, M., 2005. *Characterising Information Acquisition in Engineering Design*. PhD Thesis. Cambridge University.

BACH, K. and HARNISH, R., M, 1979. *Linguistic Communication and Speech Acts*, Cambridge, USA: MIT Press.

BACKHOUSE, C. J., BROOKES, N. J. and THE DESIGN COUNCIL, 1996. *Concurrent Engineering: What's Working Where?*, Gower Publishing LTD.

BADKE-SCHAUB, P. and FRANKENBERGER, E., 1999. Analysis of Design Projects, *Design Studies*, 20 (5) 465-480.

BAIRD, F., MOORE, C. J. and JAGODZINSKI, A. P., 2000. An ethnographic study of engineering design teams at Rolls-Royce Aerospace *Design Studies*, 21 (4) 333-355.

BALES, R. F., 1950. A Set of Categories for the Analysis of Small Group Interaction, *American Sociological Review*, 15 (2) 7.

BALES, R. F. S., F., 1951. Phases in Group Problem Solving, *Journal of Abnormal and Social Psychology*, 46 485-495.

BALL, A., PATEL, M., MCMAHON, C. A., CULLEY, S., GREEN, S. and CLARKSON, J. P., 2006. A Grand Challenge: Immortal Information and Through-Life Knowledge Management (KIM), *International Journal of Digital Curation*, 1 (1) 7.

BALOG, K. and RIJKE, M. D., 2006. Finding experts and their Details in e-mail corpora. *Proceedings of the 15th international conference on World Wide Web*, Edinburgh, Scotland. ACM,

BARTLETT, C. A. and GHOSAL, S., 1990. Matrix Management: not a structure, a frame of mind, *Harvard Business Review*, 68 138-145.

BAXTER, D., GAO, J., CASE, K., HARDING, J., YOUNG, B., SEAN, C. and DANI, S., 2007. An Engineering Design Knowledge Reuse Methodology Using Process Modelling, *Research in Engineering Design*, 18 (1).

BELBIN, R. M., 1981. *Management teams : why they succeed or fail*, Oxford.: Butterworth-Heinemann.

BELLOTTI, V., DUCHENEAUT, N., HOWARD, M., SMITH, I. and GRINTER, R. E., 2005. Quality Versus Quantity: E-Mail-Centric Task Management and Its Relation With Overload, *Human-Computer Interaction*, 20 (1) 89 - 138.

BELLOTTI, V., DUCHENEAUT, N., HOWARD, M. and SMITH, T., (2003) Taking email to task: The design and evaluation of a task management centered email tool. Ft. Lauderdale, FL, United States, Association for Computing Machinery.

BLANCHARD, B. S., 2004. *Systems Engineering Management*, Hoboken, NJ: John Wiley and Sons, Inc.

BLUM, B. A., 2007. *Contracts*, New York: Aspen Publishers.

BLUMBERG, R. and ATRE, S., (2003) Automatic Classification: Moving to the Mainstream  
*Information Management Magazine*. April, [online] Available from: <http://www.information-management.com/issues/20030401/6501-1.html> [accessed 1st June 2010].

BOARDMAN, R., 2004. *Improving Tool Support for Personal Information Management*. PhD Thesis. Imperial College London, University of London.

BOEING, 2009. *Dreamliner 787 Family Background* [online] Available from: <http://www.boeing.com/commercial/787family/background.html>, [accessed 9th November].

BOGERS, T. and VAN BOSCH, A. D., (2006) Authoritative re-ranking of search results. London, United Kingdom, Springer Verlag, Heidelberg, D-69121, Germany.

BOONE, G., 1998. Concept features in Re:Agent, an intelligent Email agent. *Proceedings of the second international conference on Autonomous agents*, Minneapolis, Minnesota, United States. ACM,

BOUARFA, H. and MOHAMED, A., 2010. The Management of Virtual Collaborative Design, *International Journal of Networking and Virtual Organisations*, 7 (1) 99 - 107.

BRACEWELL, R., GOURTOVAIA, M., WALLACE, K. and CLARKSON, J., 2007. Extending Design Rationale To Capture A Integrated Design Information Space. *International Conference On Engineering Design*, Paris, France.

BRACEWELL, R., WALLACE, K., MOSS, M. and KNOTT, D., 2009. Capturing design rationale *Computer-Aided Design*, 41 (3) pages 173-186.

BRACEWELL, R. H. and WALLACE, K., M., 2003. A Tool For Capturing Design Rationale. *International Conference On Engineering Design ICED 03*, August 19-21, 2003 Stockholm.

BRERETON, M. F., CANNON, D. M., MADBOGUNJE, A. and LEIFER, L. J., 1996. Collaboration in Design Teams. In N. Cross, H. Christinaans and K. Doest (eds.) *Analyzing Design Activity*. Wiley & Sons.

BRITISH CHAMBER OF COMMERCE THAILAND, 2009. *Companies Directory - Triumph Motorcycles (Thailand) Limited* [online] Available from: <http://members.bccthai.com/BCCT/asp/corpdetail.asp?CorpID=46>, [accessed 16/11/09].

BROWN, B. A. T., SELLEN, A. J. and O'HARA, K. P., 2000. A diary study of information capture in working life. *SIGCHI Conference on Human Factors in Computing Systems*, The Hague, The Netherlands. 438 - 445

BUCKINGHAM SHUM, S., 2009. *Compendium* [online] Available from: <http://compendium.open.ac.uk/institute/>, [accessed 1st June 2010].

BUEDE, D. M., 2009. *The Engineering Design Of Systems: Models and Methods*, 2nd edition, New Jersey: John Wiley & Sons Inc.

BUSBY, J. S., 1999. The Problem with Design Reuse: An Investigation into Outcomes and Antecedents, *Journal of Engineering Design*, 10 (3).

CAIRNCROSS, F., 1997. *The Death of Distance: How the Communications Revolution Will Change Our Lives*, Harvard Business School Press.

CAMPBELL, C. S., MAGLIO, P. P., COZZI, A. and DOM, B., (2003) Expertise identification using email communications. New Orleans, LA, United States, Association for Computing Machinery, New York, NY 10036-5701, United States.

CARMICHAEL, D. G., 2000. *Contracts and International Project Management*, Rotterdam: A. A. Balkema.

CARVER, L., THORLEY, P. and BRIDGES, K., (2006) Dealing with large amounts of information, Airbus Research Review. Bremen.

CASTELLS, M., 2000. Materials for an Exploratory Theory of the Network Society, *British Journal of Sociology*, 51 (1) 5-24.

CASTELLS, M., 2010. *The Rise Of The Network Society, The Information Age: Economy Society and Culture*, 2nd edition, Chichester: Wiley-Blackwell.

CAVES, R., 2000. *Creative Industries*, Cambridge, MA: Harvard University Press.

CAYZER, S., 2004. Semantic blogging and decentralized knowledge management, *Communications of the ACM*, 47 (12) 47-52.

CHESBROUGH, H., WILLIAM, 2006. *Open business models : the new imperative for creating and profiting from technology*, Boston, Massachusetts, USA.: Harvard Business School.

CHEUNG, W. M., NEWNES, L., MILEHAM, A. R., MARSH, R. and LANHAM, J. D., 2007. A Study of Life Cycle Costing in the Perspectives of Research and Commercial Applications in the 21st Century. *ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Design Conference*, Las Vegas, Nevada, USA.

CLARKSON, J. P. and ECKERT, C., 2005. *Design Process Improvement: A review of current practice*, Springer Verlag.

COHEN, J., 1960. A coefficient of agreement for nominal scales, , *Educational and Psychological Measurement*, 20 (1) 37-46.

COHEN, W. W., CARVALHO, V. R. and MITCHELL, T. M., 2004. Learning to Classify Email into "Speech Acts". *Empirical Methods in Natural Language Processing*, July 25-26 Barcelona, Spain.

COLE, R., EKLUND, P. and STUMME, G., 2003. Document Retrieval For E-Mail Search And Discovery Using Formal Concept Analysis, *Applied Artificial Intelligence*, 17 (3) 257-280.

CONKLIN, J., 2006. *Dialogue Mapping*, Chichester, England: John Wiley.

CONSTANT, D., KIESLER, S. and SPROULL, L., 1994. What's Mine Is Ours, or Is It? A Study of Attitudes about Information Sharing, *Information Systems Research*, 5 (4) pages 400-421.

CONWAY, A. P., WODEHOUSE, A., ION, W. J. and P, J. N., 2007. A Study of Information and Knowledge Generate During Engineering Design Meetings. *International Conference on Engineering Design ICED 07*, 28th-31st August Paris, France.

CORSTON-OLIVER, S., RINGGER, E., GAMON, M. and CAMPBELL, R., 2004. Task-focused Summarization of Email. In M.-F. Moens and S. Szpakowicz (eds.) *Text Summarization Branches Out: Proceedings of the ACL-04 Workshop*, Barcelona, Spain. Association for Computational Linguistics, 43--50.

COURT, A. W., ULLMAN, D G AND CULLEY, S J, 1998. A comparison between the provision of information to engineering designers in the UK and the USA. *Int. J. of Information Management*, 18 (6) 409-425.

CRAWFORD, E., KAY, J. and MCCREATH, E., 2002. An intelligent interface for sorting electronic mail. *Proceedings of the 7th international conference on Intelligent user interfaces*, San Francisco, California, USA. ACM,

CRISTIANINI, N., 2000. *An introduction to support vector machines : and other kernel-based learning methods*, Cambridge: Cambridge University Press.

CROSS, N. and CLAYBURN CROSS, A., 1995. Observations of Teamwork and Social Processes in Design, *Design Studies*, 16 (2) 145-170.

CROSS, N., DORST, K. and ROOZENBURG, N., 1992. *Research in Design Thinking*, Delft: Delft University Press.

CROSS, R., PARKER, A., PRUSAK, L. and BORGATTI, S. P., 2001. Knowing What We Know: Supporting Knowledge Creation and Sharing in Social Networks, *Organizational Dynamics*, 30 (2) 100-120.

DAFT, R. L. and LENGEL, R. H., 1986. Organizational Information Requirements, Media Richness and Structural Design, *Management Science*, 32 (5) 554-571.

DANDY, G., WALKER, D., DANIELL, T. and WARNER, R., 2008. *Planning and design of engineering systems*, 2nd edition, New York: Taylor and Francis.

DARLINGTON, M. J., 2002. *Cognition and the Engineering Design Requirement*. PhD Thesis. University of Bath.

DAVIES, A., BRADY, T. and TANG, P., 2003. Delivering Integrated Solutions. SPRU/CENTRIM Brighton

DAY, E., 2009. How the tyranny of email keeps us in touch - but drives us all apart. *The Observer: Focus*, 1st September, Available from [accessed

DELPHI, 2008. *Information intelligence* [online] Available from: <http://www.information-age.com/articles/283621/information-intelligence.html>, [accessed 10th March 2008].

DEPARTMENT FOR CULTURE MEDIA AND SPORT AND DEPARTMENT FOR BUSINESS INNOVATION AND SKILLS, 2009. Digital Britain: Final Report. The Stationary Office London.

DESTEFANI, J., (2004) A Look at Boeing's Outsourcing Strategy. *Manufacturing Engineering*. [online] Available from: [http://findarticles.com/p/articles/mi\\_qa3618/is\\_200403/ai\\_n9375403/pg\\_4/?tag=content;col1](http://findarticles.com/p/articles/mi_qa3618/is_200403/ai_n9375403/pg_4/?tag=content;col1) [last accessed 23/12/09], ProQuest Information and Learning Company.

DIGITAL CURATION CENTRE, 2009. *Digital Curation Life Cycle Model* [online] Available from: <http://www.dcc.ac.uk/docs/publications/DCCLifecycle.pdf>, [accessed 20th June 2009].

DOM, B., EIRON, I., COZZI, A. and ZHANG, Y., (2003) Graph-Based Ranking Algorithms for E-mail Expertise Analysis. *8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*.

DOMINICH, S., 2008. *The modern algebra of information retrieval*, Berlin: Springer-Verlag

DONATH, J., 2002. A semantic approach to visualizing online conversations, *Communications of the ACM*, 45 (4) 45-49.

DONG, A., 2005. The latent semantic approach to studying design team communication, *Design Studies*, 26 (5) 445-461.

DONG, A. and VANDE MOERE, A., (2006) Computationally assessing conceptual coherence in teamwork. *IEA2006 Congress* Maastricht, Netherlands:, Elsevier, Ltd.

DORST, K. and CROSS, N., 2001. Creativity in the design process: co-evolution of problem-solution, *Design Studies*, 22 (5) 425-437.



DUCHENEAUT, N. and BELLOTTI, V., 2001. Email as habitat: An exploration of embedded personal information management, *Interactions*, 8 30-38.

ECKERT, C., 2001. The Communication Bottleneck in Knitwear Design: Analysis and Computing Solutions, *Computer Supported Cooperative Work (CSCW)*, 10 29-74.

ECKERT, C. and BOUJUT, J.-F., 2003. The Role of Objects in Design Co-Operation: Communication through Physical or Virtual Objects, *Computer Supported Cooperative Work (CSCW)*, 12 (2) 145-151.

ECKERT, C., CLARKSON, J. and STACEY, M., (2001) Information Flow In Engineering Companies: Problems and Their Causes. IN S. Culley (Ed.) *International Conference on Engineering Design (ICED)*. Glasgow.

ECKERT, C., JOWERS, I. and CLARKSON, J. P., 2007. Knowledge Requirements Over Long Product Life Cycles. *KIM Project Conference 2007*, 28 - 29 March 2007 Loughborough University, UK.

ECKERT, C., MAIER, A. and MCMAHON, C., 2005. Communication in Design. In Eckert C and J. P. Clarkson (eds.) *Design Process Improvement: A Review of Current Practice*. Springer Verlag.

ECKERT, C. M. and STACEY, M. K., (2001) Dimensions of communication in design. *13th International Conference on Engineering Design (ICED'01)*. Glasgow, Scotland, UK.

EPPLER, M. J. and MENGIS, J., 2004. The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines, *The Information Society*, 20 (5) 325-344.

ERICSSON, K. A. and SIMON, H. A., 1993. *Protocol analysis: verbal reports as data*, revised, Cambridge, MA: MIT Press.

FAULCONBRIDGE, I. and RYAN, M. J., 2002. *Managing complex technical projects: a systems engineering approach*, Massachusetts: Artech House.

FIELD, M. and KELLER, L., 1998. *Project Management*, International Thompson Series Press.

FINANCIAL TIMES, 2010. Mobile phones are harnessed to the wisdom of crowds. *Financial Times*, 28th April 2010, Available from <http://www.ft.com/cms/s/0/957208bc-52e1-11df-813e-00144feab49a.html>, [accessed 1st June 2010].

FINCH, J., 2007. Watchdog intensifies supermarket inquiry. *The Guardian*, 20th August, Available from

<http://www.guardian.co.uk/business/2007/aug/20/supermarkets>,  
[accessed 20th October 2008].

FISHER, D., BRUSH, A. J., GLEAVE, E. and SMITH, M., A. , 2006. Revisiting Whittaker \& Sidner's "email overload" ten years later. *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, Banff, Alberta, Canada. ACM,

FLEW, T., 2002. Beyond *ad hoc*ery. Defining Creative Industries *The Second International Conference on Cultural Policy Research*, 23rd -26th January Te Papa, Wellington, New Zealand.

FORSBERG, K. and MOOZ, H., 1992. The Relationship of Systems Engineering to the Project Cycle, *Engineering Management Journal*, 4 (3) 36-43.

FRAU, S., ROBERTS, J., C. and BOUKHELIFA, N., 2005. Dynamic Coordinated Email Visualization. In V. Skala (ed.) *WSCG05 - 13th International Conference on Computer Graphics, Visualization and Computer Vision'2005*, Jan 31 - Feb 4 Plzen, Czech Republic. 187-193.

FRENCH, M., J, 1985. *Conceptual Design for Engineers*, 2nd edition, New York, USA: Springer-Verlag.

FRIEDMAN, T. L., 2005. *The World Is Flat: A Brief History of the Twenty-first Century* 3rd edition, New York: Farrar, Straus and Giroux.

FULK, J. and DESANCTIS, G., 1995. Electronic Communication and Changing Organizational Forms, *Organization Science*, 6 (4) 337-349.

GALLIERS, R. D., LEIDNER, D. E. and BAKER, B. S. H., 1999. *Strategic Information Management: Challenges and Strategies Managing Information Systems*, 2nd edition, Butterworth-Heinemann.

GANTZ, J., REINSEL, D., CHUTE, C., SCHLICHTING, W., MCARTHUR, J., MINTON, S., XHENETI, I., TONCHEVA, A. and MANFREDIZ, A., 2007. The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010. IDC, Massachusetts.

GANTZ, J., REINSEL, D., CHUTE, C., SCHLICHTING, W., MCARTHUR, J., MINTON, S., XHENETI, I., TONCHEVA, A. and MANFREDIZ, A., 2008. The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide Information Growth Through 2011. IDC Massachusetts.

GENBORG, I., 2005. Litigation's domain covers all email, *ENR (Engineering News-Record)*, 254 (18) 47.

GENEVIEVE, M. J., 2006. Synchronous and Asynchronous Text-Based CMC in Educational Contexts: A Review of Recent Research, *TechTrends*, 50 (4) 46-53.

GERO, J. and MC NEILL, T., 1998. An approach to the analysis of design protocols, *Design Studies*, 19 (1) 21-61.

GIBSON, C. B. and COHEN, S. G., 2003. *Virtual Teams That Work: Creating Conditions for Virtual Team Effectiveness*, San Francisco, CA: Jossey-Bass.

GIESS, M. D., GOH, Y. M., DING, L. and MCMAHON, C. A., 2007. Improved Product, Process and Rationale Representation and Information Organisation to Support Design Learning. *International Conference on Engineering Design*, Paris, France.

GOEL, V., 1995. *Sketches of thought*, Cambridge, MA: MIT Press.

GOH, Y. M. and MCMAHON, C., 2009. Improving reuse of in-service information capture and feedback, *Journal of Manufacturing Technology Management*, 20 (5) pages 626-639.

GORSE, C. A. and EMMITT, S., 2003. Investigating interpersonal communication during construction progress meetings: challenges and opportunities, *Engineering, Construction and Architectural Management*, 10 (4) 234-244.

GORSE, C. A. and EMMITT, S., 2007. Communication behaviour during management and design team meetings: a comparison of group interaction, *Construction Management and Economics*, 25 (11) 1197 - 1213.

GORTON, D., MURTHY, U., VEMURI, S. and PEREZ-QUINONES, M. A., 2007. *Email-Set Visualization: Facilitating Re-Finding in Email Archives*, Virginia: Virginia Tech.

GRAY, C. F. and LARSON, E. W., 2000. *Project Management: The Managerial Process*,

GREBICI, K., YEE MEY, G., SUNNY, Z., BLANCO, E. and MCMAHON, C., (2007) Information Maturity Approach for the Handling of Uncertainty within a Collaborative Design Team. *Computer Supported Cooperative Work in Design, 2007. CSCWD 2007. 11th International Conference on*.

GUBA, E. G., 1978. *Towards a methodology of naturalistic inquiry in educational evaluation*, Los Angeles: UCLA Center for the Study of Evaluation.

GUNENDRAN, (2007) Organising structured & Unstructured Information for Long Term Retrievability. *KIM Conference*.

GWIZDKA, J., (2002) Reinventing the inbox - Supporting the management of pending tasks in email. Minneapolis, MN, United States, Association for Computing Machinery.

GWIZDKA, J. and CHIGNELL, M., 2004. Individual differences and task-based user interface evaluation: A case study of pending tasks in email, *Interacting with Computers*, 16 (4) 769-797.

HABERBERG, A. and RIEPLE, A., 2001. *The Strategic Management of Organisations*, Pearson Education Limited.

HALES, C. and GOOCH, S., 2004. *Managing Engineering Design*, 2nd edition, London: Springer.

HARTLEY, J. (ed.) 2005 *Creative Industries*, Blackwell Publishing.

HEIMAN, B., NICKERSON, J. and ZENGER, T., 2009. Governing Knowledge Creation: a problem-finding and problem-solving perspective. In Nicolai J. Foss and Snejina Michailova (eds.) *Knowledge Governance: Processes and Perspectives*. Oxford, Oxford University Press.

HERTZUM, M. and PEJTERSEN, A. M., 2000. The information-seeking practices of engineers: searching for documents as well as for people, *Information Processing & Management*, 36 (5) 761-778.

HICKS, B., DONG, A., R, P. and MCALPINE, H., 2008. Organising and Managing personal electronic files: a mechanical engineer's perspective, *ACM Transactions on Information Systems* 26 (4) 1 - 40.

HICKS, B. J., CULLEY, S. J., ALLEN, R. D. and MULLINEUX, G., 2002. A framework for the requirements of capturing, storing and reusing information and knowledge in engineering design, *International Journal of Information Management*, 22 (4) 263-280.

HIGA, K., SHENG, O. R. L., BONGSIK, S. and FIGUEREDO, A. J., 2000. Understanding relationships among teleworkers' e-mail usage, e-mail richness perceptions, and e-mail productivity perceptions under a software engineering environment, *Engineering Management, IEEE Transactions on*, 47 (2) 163-173.

HILTZ, S. R., JOHNSON, K. and RABKE, A. M., (1980) The process of communication in face to face vs. computerized conferences: a controlled experiment using Bales Interaction Process Analysis. *Proceedings of the 18th annual meeting on Association for Computational Linguistics*. Philadelphia, Pennsylvania, Association for Computational Linguistics Morristown, NJ, USA.

HINDUS, D. and SCHMANDT, C., 1992. Ubiquitous audio: Capturing spontaneous collaboration. *ACM Computer Supported Cooperative Work*, Toronto, ON, Canada. 210-217.

HONEYCUTT, L., 2001. Comparing E-Mail and Synchronous Conferencing in Online Peer Response, *Written Communication*, 18 (1) 26-60.

HUBKA, V., 1988. *Practical studies in systematic design*, London: Butterworth.

HUBKA, V. and EDER, W. E., 1982. *Principles of engineering design*, London: Butterworth Scientific.

HUET, G., 2006. *Design Transaction Monitoring: Understanding Design Reviews For Extended Knowledge Capture*. PhD Thesis. University of Bath.

IBM, 2008. *Lotus Software Suite Overview, Webpage* [online] Available from: <http://www-306.ibm.com/software/uk/lotus/>, [accessed 12th March 2008].

IBM, 2009. *Information lifecycle management solutions: Optimizing storage efficiency* [online] Available from: <http://www-03.ibm.com/systems/storage/solutions/ilm/>, [accessed 22nd June 2009].

IJSSELSTEIJN, W., 2003. Staying in Touch. Social Presence and Connectedness through Synchronous and Asynchronous Communication Media. *Proceedings of HCI International Conference on Human-Computer Interaction*, New Jersey Lawrence Erlbaum Associates, 924-928.

JACKSON, T., W. , DAWSON, R. and WILSON, D., 2003. Understanding email interaction increases organizational productivity, *Commun. ACM*, 46 (8) 80-84.

JACKSON, T. W. and BURGESS, A., 2003. Capturing and Managing Email Knowledge. *Business Innovation in the Knowledge Economy - Abstracts from the IBM & Stratford-Upon-Avon Conference*, De Montfort University, Leicester, UK; Warwick & Stratford-Upon-Avon.

JACKSON, T. W., BURGESS, A. and EDWARDS, J., 2006. A simple approach to improving email communication, *Communications of the ACM*, 49 (6) 107-109.

JARVENPAA, S. L. and LEIDNER, D. E., 1999. Communication and Trust in Global Virtual Teams, *Organization Science*, 10 (6) 791-815.

KAHN, K. B., 2005. *PDMA Handbook of New Product Development* 2nd edition, John Wiley & Sons

KAPTELININ, V., (2003) UMEA: Translating interaction histories into project contexts. Ft. Lauderdale, FL, United States, Association for Computing Machinery.

KATO, Y., TAKETA, K. and HORI, K., (2002) Capturing Design Rationale by Annotating E-mails. *6th World Multiconference on Systematics, Cybernetics and Informatics*. Orlando.

KENDRICK, T., 2003. *Project Risk: Essential tools for failure-proofing your project*, New York: American Management Association.

KERZNER, H., 2006. *Project Management: A systems approach to planning, scheduling and controlling*, Hoboken, NJ: John Wiley and Sons Inc.

KHADILKAR, D. V. and STAUFFER, L. A., 1996. An experimental evaluation of design information reuse during conceptual design, *Journal of Engineering Design*, 7 (4).

KIM, S., 2002. *User modelling for knowledge sharing in e-mail communication*. PhD Thesis. Southampton.

KLEINSMANN, M. and VALKENBURG, R., 2008. Barriers and enablers for creating shared understanding in co-design projects, *Design Studies*, 29 (4) 369-386.

KLIMT, B. and YANG, Y., (2004) The enron corpus: A new dataset for email classification research. Pisa, Italy, Springer Verlag, Heidelberg, D-69121, Germany.

KOPRINSKA, I., POON, J., CLARK, J. and CHAN, J., 2007. Learning to classify e-mail, *Information Sciences*, 177 (10) 2167-2187.

KRIPPENDORFF, K., 1980. *Content Analysis: An introduction to its methodology*, Beverly Hills & London: Sage Publications Inc.

KUNZ, W. and RITTEL, H. W. J., (1970) Issues as Elements of Information Systems. Working Paper 131. Centre for Planning and Development Research.

KURASAKI, K. S., 2000. Intercoder Reliability for Validating Conclusions Drawn from Open-Ended Interview Data, *Field Methods*, 12 (3) 179-194.

KWASITSU, L., 2003. Information-seeking behavior of design, process, and manufacturing engineers, *Library & Information Science Research*, 25 (4) 459-476.

LEE, A. S., 1994. Electronic Mail as a Medium for Rich Communication: An Empirical Investigation Using Hermeneutic Interpretation, *MIS Quarterly*, 18 (2) 143-157.

LEINER, B. M., CERF, V. G., CLARK, D. D., KAHN, R. E., LEONARD, K., LYNCH, D. C., POSTEL, J., ROBERTS, L. G. and WOLFF, S. S., 1997. The past and future history of the Internet, *Communications of the ACM*, 40 (2) 102-108.

LI, Z. and RAMANI, K., 2007. Ontology-based design information extraction and retrieval, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, 21 (2) 137-154.

LIANGA, T.-P., YANGB, Y.-F., CHENC, D.-N. and KUD, Y.-C., 2008. A semantic-expansion approach to personalized knowledge recommendation, *Decision Support Systems*, 45 (3) 401-412.

LICHTENSTEIN, S., (2004) Knowledge development and creation in email. Big Island, HI., United States, Institute of Electrical and Electronics Engineers Computer Society, Piscataway, NJ 08855-1331, United States.

LIGHTMAN, E. J., MCCARTHY, P. M., DUFTY, D. F. and MCNAMARA, D. S., (2007) Cohesion and Structural Organization in High School Texts. *The Florida Artificial Intelligence Research Society International Conference (FLAIRS)*. Menlo Park, CA, AAAI Press.

LINCOLN, Y. S. and GUBA, E. G., 1985. *Naturalistic Inquiry*, New Deli & London: Sage Publications Inc. .

LOFTUS, C., HICKS, B. and MCMAHON, C., 2009. Capturing key relationships and stakeholders over the product lifecycle: an email based approach. *6th International Conference on Project LifeCycle Management (PLM 09)*, 6-8 July 2009 Bath UK.

LOMBARD, M., SNYDER-DUCH, J. and BRACKEN C. C., 2002. Content Analysis in Mass Communication: Assessment and Reporting of Inter-coder Reliability, *Human Communication Research*, 28 (4) 587-604.

LOWE, A., 2002. *Studies of information use by engineering designers and the development of strategies to aid in its classification and retrieval*. PhD Thesis. University of Bristol.

LOWE, A., MCMAHON, C., SHAH, T. and CULLEY, S., (1999) A Method for the Study of Information Use Profiles for Design Engineers. *ASME DETC Las Vegas*, Nevada.

LUSK, E. J., 2006. Email: Its decision support systems inroads-An update, *Decision Support Systems*, 42 (1) 328-332.

MACKAY, W. E., (1988) More than just a communication system: Diversity in the use of electronic mail. *Proceedings of the CSCW 89 Conference on Computer Supported Cooperative Work*. Portland, Oregon, USA, New York: ACM.

MALONE, T. W., GRANT, K. R., TURBAK, F. A., BROBST, S. A. and COHEN, M. D., 1987. Intelligent information-sharing systems, *Communications of the ACM*, 30 (5) 390-402.

MARKOPOULOS, P., IJSELSTEIJN, W., HUIJNEN, C., ROMIJN, O. and PHILOPOULOS, A., 2003. Supporting social presence through asynchronous awareness systems. In F. Davide, G. Riva and W. A. Ijsselsteijn (eds.) *Being There - Concepts, Effects and Measurement of User Presence in Synthetic Environments*. Amsterdam, IOS Press.

MARKUS, M. L., 2001. Toward a theory of knowledge reuse: types of knowledge reuse situations and factors in reuse success, *Journal of Management Information Systems*, 18 (1) 57-93.

MARSLAND, S., 2009. *Introduction to Machine Learning: An algorithmic perspective*, Chapman & Hall.

MASSEY, A. P., MONTOYA-WEISS, M. M. and YU-TING, H., 2003. Because Time Matters: Temporal Coordination in Global Virtual Project Teams, *Journal of Management Information Systems*, 19 (4) 129-155.

MAZMANIAN, M., YATES, J. and ORLIKOWSKI, W., 2006. Ubiquitous Email: Individual Experiences And Organizational Consequences of Blackberry Use. *65th Annual Meeting of the Academy of Management*, , Atlanta.

MCALPINE, H., 2007. The Investigation Of An Integrated E-Logbook To Support Information Reuse In The Engineering Domain. Department of Mechanical Engineering, University of Bath.

MCALPINE, H., 2010. *Improving the Management of Informal Engineering Information Through Electronic Logbooks*. PhD Thesis. University of Bath.

MCALPINE, H., HICKS, B. and CULLEY, S., 2006. Attempting to Investigate the Design Information loss Between Logbooks and Formal Engineering Reports. In D. R. P. a. P. Workshop (ed.) *2nd Conference on Design Computing and Cognition (DCC '06)*, 9 -13th July 2006 Eindhoven, The Netherlands.

MCCREATH, E. and KAY, J., 2003. IEMS: Helping Users Manage Email. *User Modelling 2003*.

MCDOWELL, L., ETZIONI, O. and HALEVY, A., 2004. Semantic email: Theory and applications, *Web Semantics*, 2 (2) 153-183.

MCMAHON, C., LOWE, A. and CULLEY, S., 2004. Knowledge management in engineering design: personalization and codification, *Journal of Engineering Design*, 15 (4) 307-325.

MCMAHON, C., LOWE, A., CULLEY, S., CORDEROY, M., CROSSLAND, R., SHAH, T. and STEWART, D., 2004. Waypoint: An Integrated Search and Retrieval System for Engineering Documents, *Journal of Computing and Information Science in Engineering*, 4 (4) 329-338.

MCMAHON, C. A., GIESS, M. D. and CULLEY, S., 2005. Information management for through-life product support: the curation of digital engineering data, *International Journal of Product Lifecycle Management* 1(1) 26-42.



- MCMANUS, D. J., SANKAR, C. S., CARR, H. H. and FORD, F. N., 2002. Intraorganizational Versus Interorganizational Uses and Benefits of Electronic Mail, *Information Resources Management Journal*, 15 (3).
- MEADOW, C. T., BOYCE, B. R., KRAFT, D. H. and BARRY, C., 2007. *Text Information Retrieval Systems*, 3rd edition, London: Academia Press.
- MEDLAND, A. J., 1992. *The computer-based design process*, 2nd edition, London: Chapman & Hall.
- MEDLAND, A. J., 1992. Forms of Communications Observed During the Study of Design Activities in Industry, *Journal of Engineering Design*, 5 243-253.
- MINISTRY OF DEFENCE, 2009. *The CADMID Lifecycle* [online] Available from: <http://www.aof.mod.uk/aofcontent/tactical/ppm/content/lifecycles/cadmid.htm>, [accessed 26th January 2010].
- MINKOV, E., COHEN, W. W. and NG, A. Y., (2006) Contextual search and name disambiguation in email using graphs. Seattle, WA, United States, Association for Computing Machinery, New York, NY 10036-5701, United States.
- NELSON, L. and CHURCHILL, E. F., 2007. Mapmail: restructuring an email client for use in distributed teams. *CHI '07 extended abstracts on Human factors in computing systems*, San Jose, CA, USA. ACM Press,
- NENKOVA, A. and BAGGA, A., (2003) Email classification for contact centers. Melbourne, FL, United States, Association for Computing Machinery.
- NEUSTAEDTER, C., BRUSH, A. J. B. and SMITH, M., A., 2005. Beyond "from" and "received": exploring the dynamics of email triage. *CHI '05 extended abstracts on Human factors in computing systems*, Portland, OR, USA. ACM,
- NEWELL, A. and SIMON, H. A., 1972. *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall.
- NEWNES, L., MEDLAND, A., DANIEL, J. and HOLMES, C., (2006) A Cyclic Design Approach Applied to a Cervical Screening Device. *Proceedings of the Institution of Mechanical Engineers -- Part B -- Engineering Manufacture*. Professional Engineering Publishing.
- NEWNES, L., MILEHAM, A. R., CHEUNG, W. M., MARSH, R., LANHAM, J. D., SARAVI, M. E. and BRADBURY, R. W., 2008. Predicting the Whole Life Cost of a Product at the Conceptual Design Stage, *Journal of Engineering Design*, 19 (12) 99 - 112.

NORRIS, G., 2005. *Airbus A380: Super jumbo of the 21st century*, Minnesota: MBI Publishing Company.

O'KANE, P. and HARGIE, O., 2007. Intentional and unintentional consequences of substituting face-to-face interaction with e-mail: An employee-based perspective, *Interacting with Computers*, 19 (1) 20-31.

O'KANE, P., PALMER, M. and HARGIE, O., 2007. Workplace interactions and the polymorphic role of e-mail, *Leadership & Organization Development Journal*, 28 (4) 308-324.

OLIVA, R. and KALLENBERG, R., 2003. Managing the transition from products to services, *International Journal of Service Industry Management*, 14 (2) 160-172.

OXMAN, R. E., 1994. Precedents in Design: A Computational Model for the Organization of Precedent Knowledge, *Design Studies*, 15 (2) 141 - 157.

OXMAN, R. E., 1996. CBR in Design, *AI Communications*, 9 (3) 117-127.

PAHL, A.-K., NEWNES, L. and MCMAHON, C., 2007. A generic model for creativity and innovation: overview for early phases of engineering design, *Journal of Design Research*, 6 (1-2) pg 5-44.

PAHL, G. and BEITZ, W., 1996. *Engineering Design: A Systematic Approach*, 2nd edition, Springer.

PANTELI, N., 2002. Richness, power cues and email text, *Information and Management*, 40 (2) 75-86.

PATTON, M. Q., 2002. *Qualitative Research and Evaluation Methods*, 3rd edition, New Delhi & London: Sage Publications Inc.

PAYNE, T. and EDWARDS, P., 1997. Interface agents that learn: An investigation of learning issues in a mail agent interface. , *Applied Artificial Intelligence*, 1-32.

PEDLEY, P., 2006. *Essential Law for Information Professionals*, London: Facet Publishing.

PENNOCK, M., 2006. "Curating E-Mails: A life cycle approach to the management and preservation of e-mail messages" *DCC Digital Curation Manual* [online] Available from: <http://www.dcc.ac.uk/resource/curation-manual/chapters/curating-e-mails/curating-e-mails>, [accessed 3rd July 2008].

PENNOCK, M., 2007. Digital Curation: A Life-Cycle Approach to Managing and Preserving Usable Digital Information, *Library & Archives*, 1 -.

PORTER, M., 1990. *The Competitive Advantage of Nations*, MacMillan Press LTD.

PRASAD, B., 1995. *Concurrent engineering companies. Integrated product and process organization.*, New Jersey: Prentice Hall.

PRUSAK, L., 2001. Where Did Knowledge Management Come From?, *IBM Systems Journal*, 40 (4) 1002.

PUGH, S., 1991. *Total Design: integrated methods for successful product engineering*, Workingham, UK: Addison-Wesley.

QIAN, L. and GERO, J. S., 1996. Function-behaviour-structure paths and their role in analogy-based design, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 10 289-312.

QSR, 2008. *NVivo 7 Overview* [online] Available from: [http://www.qsrinternational.com/products\\_nvivo.aspx](http://www.qsrinternational.com/products_nvivo.aspx), [accessed 7th March 2008].

REGLI, W. C. and CICIRELLO, V. A., 2000. Managing digital libraries for computer-aided design, *Computer-Aided Design*, 32 119-132.

REID, F. J. M., MALINEK, V., STOTT, C. J. T. and EVANS, J., 1996. The messaging threshold in computer-mediated communication, *Ergonomics*, 39 (8) 1017-1037.

RENAUD, K., RAMSAY, J. and HAIR, M., 2006. "You've Got E-Mail!" ... Shall I Deal With It Now? Electronic Mail From the Recipient's Perspective, *International Journal of Human-Computer Interaction*, 21 (3) 313 - 332.

RENNIE, J., D. M., 2000. ifile: An Application of Machine Learning to E-Mail Filtering. *KDD-2000 Text Mining Workshop*, August Boston, MA, USA.

RITTEL, H. W. J. and WEBBER, M. M., 1973. Dilemmas in a General Theory of Planning, *Policy Sciences*, 4 155-169.

ROBSON, M., 2002. *Problem Solving In Groups*, 3rd edition, Aldershot: Gower.

ROHALL, S., L. , GRUEN, D., MOODY, P., WATTENBERG, M., STERN, M., KERR, B., BOB, S., DAVE, K., ARMES, R. and WILCOX, E., 2004. ReMail: a reinvented email prototype. *CHI '04 extended abstracts on Human factors in computing systems*, Vienna, Austria. ACM,

RONG, Q., WEI, Z. and BINGRU, Y., (2006) Detect community structure from the enron email corpus based on link mining. Jinan, China, Institute of Electrical and Electronics Engineers Computer Society, Piscataway, NJ 08855-1331, United States.

RUSH, C. and RAJKUMAR, R., 2000. Analysis of Cost Estimating Processes Used Within a Concurrent Engineering Environment Throughout a Product Lifecycle. *International Conference on Concurrent Engineering CE 2000*, Lyon, France.

SAAKSVUORI, A. and IMMONEN, A., 2008. *Product Lifecycle Management*, 3rd edition, Berlin: Springer-Verlag.

SAMIEI, M., DILL, J. and KIRKPATRICK, A., (2004) EzMail: Using information visualization techniques to help manage email. London, United Kingdom, Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ 08855-1331, United States.

SCHÖN, D. A., 1983. *The reflective practitioner*, New York: Basic Books.

SCHWARTZ, D. G., 1999. When email meets organizational memories: addressing threats to communication in a learning organization, *International Journal of Human Computer Studies*, 51 (3) 599-614.

SEGAL, R., B. and KEPHART, J., O., 1999. MailCat: an intelligent assistant for organizing e-mail. *Proceedings of the third annual conference on Autonomous Agents*, Seattle, Washington, United States. ACM,

SEGAL, R., B. and KEPHART, J., O., 2000. Incremental Learning in SwiftFile. *Proceedings of the Seventeenth International Conference on Machine Learning*: Morgan Kaufmann Publishers Inc.,

SENGE, P. M., 2006. *The Fifth Discipline: The art and practice of the learning organisation*, 2nd edition, London: Arrow Books Ltd.

SHIGLEY, J. E., MISCHKE, C. R. and BUDYNAS, R. G., 2003. *Mechanical Engineering Design*, 7th edition, New York: McGraw-Hill Education.

SHIPMAN, F. and MCCALL, R., 1997. Integrating Different Perspectives on Design Rationale: supporting the emergence of design rationale from design communication., *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, 11 (2) 141-154.

SHIRANI, A. I., TAFTI, M. H. A. and AFFISCO, J. F., 1999. Task and technology fit: a comparison of two technologies for synchronous and asynchronous group communication, *Information and Management*, 36 (3) 139-150.

SIM, S. K. and DUFFY, A. H. B., 2003. Towards an ontology of generic engineering design activities, *Research in Engineering Design*, 14 (4) 200-223.

SIMON, H. A., 1969. *The Sciences of the Artificial*, Cambridge, MA: MIT Press.

SIMON, H. A., 1973. Structure of ill structured problems, *Artificial Intelligence*, 4 (3-4) 181-201.

SIMS, 2003. *How Much Information?* [online] Available from: <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>, [accessed 5th March 2008].

SIVALOGANATHAN, S. and SHAHIN, T., (1998) Engineering Design Conference '98: Design Reuse. Brunel University, Professional Engineering Publishing. .

SMITH, N. J., 2007. *Engineering Project Management*, 3rd edition, Blackwell Publishing.

SOW, D., EBLING, M., LEHMANN, R.-P., DAVIS, J. and BERGMAN, L., (2005) SCOUT contextually organizes user tasks. Beijing, China, Institute of Electrical and Electronics Engineers Computer Society, Piscataway, NJ 08855-1331, United States.

STARK, J., 2004. *Product Lifecycle Management: 21st Century Paradigm for Product Realisation*, London: Springer-Verlag.

STEMPFLE, J. and BADKE-SCHAUB, P., 2002. Thinking in design teams - an analysis of team communication, *Design Studies*, 23 (5) 473-496.

STRZALKOWSKI, T. (ed.) 1999 *Natural Language Information Retrieval (Text, Speech and Language Technology)* Springer.

SUROWIECKI, J., 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Doubleday; Anchor.

TALLON, P. P. and SCANNELL, R., 2007. Information Life cycle management, *Communication of the ACM Archive*, 50 (11) pages 65-69.

TEECE, D., J, 1986. Profiting from technological innovation: implications for integration, collaboration licensing and public policy, *Research Policy*, 15 285-305.

THAMHAIN, H. J., 2003. Managing innovative R&D teams, *R&D Management*, 33 (3) 297-311.

THE RADICI GROUP INC., 2008. Addressing Email Chaos: The email manager solution.

TOMIYAMA, T., 1998. Concurrent Engineering: A Successful Example for Engineering Design Research. In A. H. B. Duffy (ed.) *The Design Productivity Debate*. London, Springer-Verlag.

TÖRLIND, P., LARSSON, A., LÖFSTRAND, M. and KARLSSON, L., 2005. Towards True Collaboration in Global Design Teams? *International*

*Conference on Engineering Design, ICED 05*, August 15-18. Melbourne, Australia.

TRUONG, K. N. and HAYES, G. R., 2009. Ubiquitous Computing for Capture and Access, *Foundations and Trends in Human-Computer Interaction*, 2 (2) 95-171.

TURNER, D. A. and ROSS, K. W., 2000. Continuous media e-mail on the Internet: infrastructure inadequacies and a sender-side solution, *Network, IEEE*, 14 (4) pages 30-37.

TYLER, J. R., WILKINSON, D. M. and HUBERMAN, B. A., 2005. E-mail as spectroscopy: Automated discovery of community structure within organizations, *Information Society*, 21 (2) 133-141.

ULLMAN, D. G., 1992. *The Mechanical Design Process*, New York: McGraw-Hill.

ULLMAN, D. G., 2002. *The Mechanical Design Process*, 3rd edition, New York: McGraw-Hill.

ULRICH, D., G and EPPINGER, S., D, 1995. *Product Design and Development*, New York, USA: McGraw-Hill Inc.

VALKENBURG, R. and DORST, K., 1998. Reflective practice of design teams, *Design Studies*, 19 (3) 249-271.

VELOSO, A. and MEIRA, W., JR., (2006) Lazy associative classification for content-based spam detection. Cholula, Mexico, Institute of Electrical and Electronics Engineers Computer Society, Piscataway, NJ 08855-1331, United States.

VENOLIA, G. D., DABBISH, L., CADIZ, J. J. and GUPTA, A., 2001. Supporting Email Workflow. Microsoft Corporation.

VIEGAS, F. B., GOLDER, S. and DONATH, J., (2006) Visualizing email content: Portraying relationships from conversational histories. Montreal, QC, Canada, Association for Computing Machinery, New York, NY 10036-5701, United States.

WANG, J. X. and ROUSH, M. L., 2000. *What every engineer should know about risk engineering and management*, Basel: Taylor & Francis Inc. Marcel Dekker Inc.

WASIAK, J. O., NEWNES, L. B., MILEHAM, A. R., HICKS, B. J. and OUTRAM, G., (2008) Guiding Team Selection And The Use Of The Belbin Approach. IN D. Marjanovic, M. Storga, N. Pavkovic and N. Bojcetic (Eds.) *International Design Conference - Design 2008*. Dubrovnik, Croatia, Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb and The Design Society, Glasgow.

WATTENBERG, M., ROHALL, S. L., GRUEN, D. and KERR, B., 2005. E-Mail Research: Targeting the Enterprise, *Human-Computer Interaction*, 20 (1) 139 - 162.

WEBBER, J., (2009) Boeing to Rein in Dreamliner Outsourcing. *Business Week*. Available online [[http://www.businessweek.com/bwdaily/dnflash/content/jan2009/db20090116\\_971202.htm](http://www.businessweek.com/bwdaily/dnflash/content/jan2009/db20090116_971202.htm)], Bloomberg, New York.

WENGRAF, T., 2001. *Qualitative Research Interviewing: Biographic Narrative and Semi-structured Methods*, London: Sage Publications.

WHITTAKER, S., BELLOTTI, V. and GWIZDKA, J., 2006. Email in personal information management, *Communications of the ACM*, 49 (1) 68-73.

WHITTAKER, S. and SIDNER, C., (1996) Email overload: exploring personal information management of email. *The 1996 Conference on Human Factors in Computing Systems*.

WHITTAKER, S., TUCKER, S., SWAMPILLAI, K. and LABAN, R., 2008. Design and evaluation of systems to support interaction capture and retrieval, *Personal and Ubiquitous Computing*, 12 (3) 197-221.

WILD, P. J., CULLEY, S., MCMAHON, C., DARLINGTON, M. and LIU, S., 2006. Starting to Audit Documents in the Engineering Domain. *Computational Science and Its Applications*. Berlin, Springer.

WILSON, E. V., 2002. Email winners and losers, *Communications of the ACM*, 45 (10) 121-126.

YANG, M. C., WOOD III, W. H. and CUTKOSKY, M. R., 2005. Design information retrieval: A thesauri-based approach for reuse of informal design information, *Engineering with Computers*, 21 (2) 177-192.

ZEID, I., 1991. *CAD/Cam Theory and Practice*, McGraw-Hill Higher Education.

ZHAO, Y., TANG, L. C. M., DARLINGTON, M. J., AUSTIN, S. A. and CULLEY, S. J., 2007. Establishing Information Valuing Characteristics for Engineering Design Information. *International Conference on Engineering Design, ICED '07*, Paris, France.

## **Appendix A Semi-Structured Interview Model**

### **Interview Plan**

The aims of the interview sessions are:

- 1 To understand the technical competency and usage of communication tools by Engineers & between Engineers within Converteam.
- 2 To validate, re-enforce and provide understanding for secondary analysis to the findings of the email analysis.
- 3 To identify what opportunities exist for improving email practices within Converteam – in addition and support of the findings from the email analysis

To achieve this five sections of semi-structured questions are presented. Questions are provided as guidelines and are suitably worded so as not to lead the interview, more than is necessary to obtain relevant responses; they are proposed at three levels. The level one questions are intended to be quite open and holistic, the level two and level three questions are designed to act as prompts if required, to help interviewees answer the broader first question. Before the questions are presented, a proposed introductory spiel is given.

### **Introductory Text**

“I’m from the IdMRC research centre at the University of Bath and work in the Design Information and Knowledge group.

Our research investigates ways of better supporting Engineers in areas such as IT, knowledge sharing and access to information resources.

For the last eight months I have been working on this project, investigating email use within Converteam. Our aims and interests from this project are both academic and practical. We're interested in looking at the longer term / bigger picture for improving information management, and to develop some more specific solutions to help Converteam. – So I hope giving your time for this interview today will be rewarded in the future with benefits we might offer.



The purpose of these interviews is to gauge how workers at Converteam use and regard email as a communication tool. Although we are interested in how different peoples' accounts may vary – we are not here to critically examine your email / work skills! Hence what is said in this interview room will not be directly passed back to anyone within the company, only the findings from our project as a whole.

If it is acceptable with you I would like to tape record the interview, as an aid for my memory afterwards, so that I can concentrate on chatting with you, and not trying to write notes at light speed. The recordings will be treated in confidence by the research team and will not be used without your expressed permission. Is that ok with you?

The tape recorder is now on, if you would like me at anytime to stop it then please indicate so.”

**Develop an understanding of how the interviewee uses email.**

Would you give me a rough outline of how your day is spent dealing with email, and what kind of things you are using it for?

*How much of your time do you think you spend working with emails?*

*How effectively do you feel that time with email is being spent?*

*Consider the time you spend using emails. What kind of activities are you undertaking and roughly what proportion of your time do you spend on each?*

Do you spend more time...      sending emails (from scratch)?  
   replying to emails?  
   reading emails?

Who are you usually making contact with when you use email?

*Do you email people outside of the company?*

*Do you spend much time emailing other colleagues in the office with you?*

*Do you communicate much with engineers from other Converteam sites?*

*Do you receive email from different groups of people from those you send email to?*

Presumably you use methods of communication other than email. Are you using them in similar or different ways and with similar or different people to when using email?

*For example do you use phone, fax, meetings, conversation in the office, instant messenger?*

*How much of your time do you spend using other communication tools?*

*How effectively do you feel that time is being spent?*

*Would you describe the tasks you use the phone/fax for?*

*Why not use email instead?*

*Why is the phone/ fax preferable to email?*

*What benefits do these other tools have compared with email?*

*Who are you usually making contact with when you use phone/fax?*

*Are the people you phone/fax/chat with the same people who you regularly email?*

*Do you call people outside of the company?*

*Do you spend much time calling other colleagues in the office with you?*

*Do you communicate much with engineers from other Converteam sites?*

## **Their Involvement in the Case Study Project**

Would you describe the role you played in the project?

*Tell me about the types of activities you were involved in over the project duration?*

*Did these change with time?*

*What managerial / supervisory responsibilities did you hold?*

Would you tell me who you communicated with during the project?

*Who did you communicate most with by email during the project?*

*Who did you communicate with externally to Converteam?*

*Who did you communicate with at other Converteam sites?*

*Were the people you emailed the same people you communicated with using other mediums?*

## **Develop an understanding of how the project group used email.**

"I would like to talk a little bit about email in relation to the [case study] project, but if you would like to be more general and include other projects, then please do so.

I want to try and establish what “role” you feel that the emails you send and receive play. By this I mean what kind of topics they relate to and what kind of activities they help achieve. I’m also interested in how this might have changed over the time frame of the project.”

What tasks do you use email for?

*Can you think of some examples for me?*

What activities do you use email for?

What activities does email help you to achieve?

What processes do you use email for / to help you with?

Would you think of some reasons that might prompt you to send and email?

Would you think of some reasons for which you might receive an email?

What topics do your emails typically relate to?

What subject are your emails usually about?

Design Information  
Management Activities  
Company Activities  
Project Related Activities  
Discussion of Ideas

What kind of role did email take in sharing / discussing design information? Rather than more general information.

*What kind of discussion about aspects of the design might have taken place?*

*What kind of things did you not use email for?*

Did you feel that email was effective for transfer of design information?

### **Develop an understanding of how the organisation uses email.**

What do you feel are the purposes and functions of, email, other electronic working tools used by Converteam?

*Do you feel that their roles are well defined?*

*Do you feel that the company has practices / protocol for using these different tools?*

*How well do email, PROMIS and other e-tools integrate?*

**Capture any thoughts and suggestions they may have**

With regards to email and any other work with electronic documents, what have you learned from other companies that you have worked for or worked with?

Do you have any comments you would like to make of Converteam's information management practices?

Do you feel that email aids or hinders your work?

*What problems do you think there are?*

*What potential do you think there is for improvement?*

*How worthwhile do you feel work into this is? – I promise not to take offense!*

And if you could change one thing about email use in general, what would it be?

*Are there other things that you would like to change?*

## **Appendix B – Five Step Guide To Email Coding**

The of coding emails was discussed in 5.5 where the final coding scheme was also presented. This section presents an expanded step by step guide for the coding of emails. It is intended to instruct would be coders in future research. Five steps to coding an email are described, following this an email is coded step by step. To conclude a further example of an email coded with What, Why and How terms is given.

### **B1 - Five Steps to Coding**

A guide to coding emails in five steps is now presented.

**Step 1. – Establish Context** Skim read through text to get the gist of the email. If it is necessary to, review any quoted text which follows the body of the message and read the subject line and email addresses. The aim is to establish the context of the message.

**Step 2. – Topic Coding, top level:** Identify the predominant top level topic code or codes; i.e. Product, Project, Company (or any combination thereof). This stage will be lead by the context of the message.

**Step 3. – Topic Coding, lower level:** Work through the text highlighting any portion of text relating to second level topic terms i.e. features, time, human resources. If terms are used which do not fall within the top level mark up step two should be returned to and reconsidered.

**Step 4 – Purpose Coding:** Work through the text highlighting any portion of text identifying an email sending purpose from the lower level codes; i.e. exploring, requesting information, managing. There is no need to code at the top level; i.e. information transaction, management transaction and problem solving contribution. These should be deduced by the lower level codes.

**Step 5 – Expression Coding:** Work through the text highlighting every portion of text as to its Bale's category. There is not a top level of coding.

## **B2 - Step by Step Coding Example**

An example of coding using the five steps is now given. The example is derived from a real message sent by the partner company. The coding scheme is first repeated.

**Table B 1: Topic Coding Scheme, used for Stage 2 and Stage 3.**

What topics does the email discuss? Subject matter it relates to		
<b>Product</b> The output of the project; it may be a physical artefact or software.	<b>Project</b> the domain within which the product is created.	<b>Company</b> the sponsors or facilitators of a project.
<b>Functions:</b> Things the product must do. e.g. Be fast	<b>Risk:</b> Assessing likelihood and weighting implications	<b>Economic Issues:</b> Costs and efficiency, market and product selection
<b>Performance:</b> How well the product achieves its functions.	<b>Plans:</b> Management of phases, activities and tasks,	<b>Human Resources:</b> People, availability, allocation, training, replacing
<b>Feature:</b> The quality or characteristic with which the function is achieved.	<b>Team:</b> Team selection, development	<b>Physical Resources:</b> Ranging from offices to equipment
<b>Operating Environment:</b> Objects that interact with the product	<b>Quality Management:</b> Quality, standard or expectations	<b>Financial Resources:</b> Cash, Assets, Borrowing.
<b>Materials and Components:</b> Materials and component selection and characteristics	<b>Cost:</b> Financial arrangements at the level of the project, rather than specific component costs.	<b>Knowledge Resources:</b> Current ability and stored information.
<b>Manufacturing:</b> Consideration of manufacturing, assembly and transport.	<b>Time:</b> Durations or deadlines. Any link or reference to time.	<b>Tools and Methods:</b> Specific testing and modelling techniques.
<b>Cost:</b> Consideration of costs particularly unit costs.	<b>Manufacture:</b> Arranging manufacture, planning manufacture, in the context of the project	<b>Practices &amp; Procedures:</b> Accumulated by the company, often developed through experience
<b>Specification:</b> Formal requirements definition for the product/design. Or requirements for sub / super components of the product.	<b>Contracts:</b> Legal arrangements involving two or more parties setting out what is required from the project, often specifying costs and time.	<b>Stake Holders:</b> Such as: share holders, customers, directors and their culture and politics. They must have more than an interest in the company but in the project.
<b>Ergonomics:</b> User Interaction with product	<b>Delivery:</b> The delivery or provision of a specific component or sub-system.	
	<b>Milestones &amp; Deliverables:</b> Targets to be achieved, or which have been achieved, related to formal stages within the project.	
	<b>Documentation or Knowledge Resources:</b> Reference to general documentation resource, most likely PROMIS or mention of Knowledge Management process specific to the project.	
	<b>Administration:</b> General administration related to the project, but not distinctly captured by one of the other terms above.	

### **Step 1. - Establish Context**

This email considered in this example is italicised below. The general gist of the email is that the sender is trying to influence the recipient to choose a composite material for the diffuser, which is to be discussed at a meeting next week.

*“I understand that you and Brian are off to Munich this week to start talking about the new diffuser. In light of this I have noted down a few points which you may want to think about when the composition of the material is being discussed. I will strongly recommend a composite for weight saving reasons, despite the increased cost. Please will you call in on my office later and we can arrange to discuss these.”*

### **Step 2. - Topic Coding, top level**

The email is initially coded for a top level topic(s), Table B 1. This email is generally about the **product**, in this case the new diffuser for a car. Although it is not critical to the coding the awareness that this is a car diffuser comes from general knowledge of the project. Although there are some references to meeting later the essence of this email is not about the project i.e. through planning or activities. The essence of the email is about the diffuser, the product.

*“I understand that you and Brian are off to Munich this week to start talking about the new diffuser. In light of this I have noted down a few points which you may want to think about when the composition of the material is being discussed. I will strongly recommend a composite for weight saving reasons, despite the increased cost. Please will you call in on my office later and we can arrange to discuss these.”*

In an alternative scenario consider an email containing only the words *“please will you call into my office later and we can arrange to discuss this.”* This email has the word diffuser in the subject line or in the quoted text it is in reply of. In this case it would still be possible to code the email as product, using the wider context gained.

### **Step 3. - Topic Coding, lower level**

The email is coded for lower level topics, Table B 1. The first code to be identified is **Materials and components**. The component is the diffuser and the material of the diffuser is also mentioned.

The second code is **performance** as it can be understood that weight saving will increase performance of this artifact. This is an occurrence of a code which would cause typical inter-coder reliability issues. The weight saving benefits have been attributed to performance. Another coder may not identify this. Both however would be likely to agree with the top level code of product. This demonstrates why the lower tier of coding is less reliable.

The third code is **cost**, again from the product group. A direct reference to increased cost to the product is made.

Other text is not specifically coded. This does not matter as the email is known to be generally about the product. There are no other concepts relating to the product which have not been coded.

The email does not tell us anymore about the *‘points which you may want to think about.’* If the meeting later was to be coded this would probably contain more information and could be coded with many other terms.



*"I understand that you and Brian are off to Munich this week to start talking about the new diffuser. In light of this I have noted down a few points which you may want to think about when the composition of the material is being discussed. I will strongly recommend a composite for weight saving reasons, despite the increased cost. Please will you call in on my office later and we can arrange to discuss these."*

**Table B 2: Purpose Coding Scheme, used for Stage 4**

Why are the emails sent? Their purpose	
Problem Solving Contribution	Information Transaction
A contribution to a problem solving process. More active than simply passing information.	A straightforward transfer of information or something enabling the transfer of information.
<b>Goal Setting:</b> Identifying where the design is, and where it needs progressing to.	<b>Informing:</b> Sharing, presenting or distribution information with others. No response is required. It is passive.
<b>Constraining:</b> Imposing boundaries with requirements and desirables	<b>Requesting Information:</b> Direct request to another party to provide information, or further information. Including explicit responses to requests for information.
<b>Exploring:</b> Discussing possibilities and ideas, invoking suggestions. A return is expected from the recipient.	<b>Clarifying:</b> Clearing up misunderstandings (both requesting and in response). Asking for explanations, resolving a general lack of clarity.
<b>Developing Solutions:</b> It may encompass one or more of the following stages: searching, gathering, creating and developing solutions. Presentation of solutions for comment is also encompassed.	<b>Management Transaction</b>
<b>Evaluating:</b> Judging the quality, value and importance of something.	An effort to achieve action by asking or telling for things to be done.
<b>Decision Making:</b> Considering key factors from evaluation and possible compromises to form decision.	<b>Managing:</b> Includes arranging, directing and instructing. Implies action (such as a response) needs to be taken. Including process management outside of the organisation, e.g. prompting arrangements / meetings with third parties.
<b>Reflecting:</b> Reflecting upon a design/product decision or process already adopted or occurred. Reflecting may question whether a new of further problem now exists.	<b>Confirming:</b> Confirming or requesting confirmation of something
<b>Debating:</b> Discussing opposite views.	

#### Step 4. – Purpose Coding

The email is coded to identify why it has been sent, Table B 2. Lower level codes are applied. The sender sets the scene by informing the recipient what he/she knows. The email carries on informing that some points have been noted down in relation to the material composition.

The text then provides an **evaluation**, choosing a composite for weight saving reasons despite cost. If this was put up for discussion i.e. ‘do you think it is worth the extra cost to use a composite for weight saving benefits?’ the code of exploring would then be used.

The final sentence is coded **managing** as it instructs the recipient to see the sender about the matter.

*“I understand that you and Brian are off to Munich this week to start talking about the new diffuser. In light of this I have noted down a few points which you may want to think about when the composition of the material is being discussed. I will strongly recommend a composite for weight saving reasons, despite the increased cost. Please will you call in on my office later and we can arrange to discuss these.”*

**Table B 3: Expression coding scheme, used for Stage 5.**

<b>How is email content expressed?</b>			
<b>Socio Emotional Terms</b>		<b>Task Related Terms</b>	
<b>Positive Reactions</b>	<b>Negative Reactions</b>	<b>Sharing</b>	<b>Requesting</b>
<i>Shows Solidarity</i> raises other's status, gives help, reward	<i>Shows Antagonism</i> deflates other's status, asserts or defends self	<i>Gives Opinion</i> , evaluation, analysis, expresses feeling or wish	<i>Asks for Opinion</i> , evaluation, analysis, expression or feeling
<i>Shows Tension Release</i> jokes, laughs, shows satisfaction	<i>Shows Tension</i> asks for help, withdraws out of field	<i>Gives Suggestion</i> direction, implying autonomy for other	<i>Asks for Suggestion</i> direction, possible ways of action.
<i>Agrees</i> , shows passive acceptance, understands, concurs, complies	<i>Disagrees</i> , shows passive rejection, formality, withholds resources	<i>Gives Orientation</i> , information, repeats, clarifies.	<i>Asks for Orientation</i> information, repetition and confirmation

### **Step 5. – Expression Coding**

The email is coded as to how the content is expressed, Table B 3. The first sentence **gives orientation** as to the forthcoming message. The

second sentence gives direction to the recipient as to what they should do, this is **giving suggestion**. The following sentence **gives suggestion** as to what the sender would do. The final sentence directs the recipient what to do, '*call in on my office*', this fits into the definition of **giving suggestion** once more.

The email is very much written in the style of a manager telling an engineer what to do, through gentle suggestion and ultimately an instruction. This is reflected in the dominance of *gives suggestion* coding.

*"I understand that you and Brian are off to Munich this week to start talking about the new diffuser. In light of this I have noted down a few points which you may want to think about when the choice of the material is being discussed. I will strongly recommend a composite for weight saving reasons, despite the increased cost. Please will you call in on my office later and we can arrange to discuss these."*

### **B3 - Further Example**

A further example of how the email coding scheme is used is now presented. This shows the final coding result rather than moving through the five steps. An email thread is quoted; this is reproduced three times with the coding for *what* topic the email relates to on the first, *why* the email is sent on the second and *how* the content is expressed on the third. For the purpose of clarity in the example, the email is broken into small paragraphs to fit the chunks of coding used. For each figure a paragraph of text presents the rationale for the coding.

The example is based on emails from the corpuses analysed for this research project, names and terminology have been changed to ensure anonymity.

#### **Example – Steve’s Retirement**

What Rationale: The email is discussing a human resource (Steve). The context is about his continued employment and wanting him in the *company* (as a *human resource*). If the email discussed keeping him on to help with a specific project then it would instead fall into the *Project* category, and under the term of *Team*. There are no other substantive topics in the email.

	<u>Specific Terms</u>	<u>General Category</u>
<p>To: Oliver Mustard From: John Green Subject: Steve's Retirement</p> <p>Oliver,</p> <p>Steve will be coming to the end of his post retirement contract in the next couple of months.</p> <p>We need to ensure he has a continuation contract to follow on as he is willing to continue and there is still plenty of work for him to do!!</p> <p>Regards, John</p>	Human Resources	Company
<p>To: John Green From: Oliver Mustard Subject: Re Steve's Retirement</p> <p>John</p> <p>Glad to hear that - both aspects.</p> <p>There does not seem to be much of a problem getting contract renewals.</p> <p>My memory is that Steve's runs till the end of March so I do not see it as being very urgent. I would expect to start looking at it in the New Year.</p> <p>Oliver</p>	Human Resources	Company

**Figure B 1: Example – Coding for What topics the email relates to**

	<u>Specific Terms</u>	<u>General Category</u>
<p>To: Oliver Mustard From: John Green Subject: Steve's Retirement</p> <p>Oliver,</p> <p>Steve will be coming to the end of his post retirement contract in the next couple of months.</p> <p>We need to ensure he has a continuation contract to follow on as he is willing to continue and there is still plenty of work for him to do!!</p> <p>Regards, John</p>	<p>Informing</p> <p>Managing</p>	<p>Information Transaction</p> <p>Management Transaction</p>
<p>To: John Green From: Oliver Mustard Subject: Re Steve's Retirement</p> <p>John</p> <p>Glad to hear that - both aspects.</p> <p>There does not seem to be much of a problem getting contract renewals.</p> <p>My memory is that Steve's runs till the end of March so I do not see it as being very urgent. I would expect to start looking at it in the New Year.</p> <p>Oliver</p>	<p>Informing</p> <p>Informing</p> <p>Informing</p>	<p>Information Transaction</p>

**Figure B 2: Example – Coding for Why the email has been sent**

Why Rationale: The first paragraph of the email on its own provides information about Steve's contract, hence it is coded *Informing*. Although the second paragraph does not instruct the recipient to do something directly it implies that action needs to be taken by that party, hence it is coded *Managing*. The responding email informs the initial sender of the recipients position, coded *informing*.

How Rationale: The opening paragraph of the first email provides some information which *orientates* the following discussion. The second then makes a *suggestion* about how to deal with Steve reaching the end of his contract. The reply begins with giving a positive *opinion*, the replier expressing their feeling on the matter. The replier then gives their *opinion* on the matter of contract renewals, this is their evaluation of the situation. Further *orientation* is added to

<p>To: Oliver Mustard From: John Green Subject: Steve's Retirement</p> <p>Oliver,</p> <p>Steve will be coming to the end of his post retirement contract in the next couple of months.</p> <p>We need to ensure he has a continuation contract to follow on as he is willing to continue and there is still plenty of work for him to do!!</p> <p>Regards, John</p>	<p><b><u>Specific Terms</u></b></p>
<p>To: John Green From: Oliver Mustard Subject: Re Steve's Retirement</p> <p>John</p> <p>Glad to hear that - both aspects.</p> <p>There does not seem to be much of a problem getting contract renewals.</p> <p>My memory is that Steve's runs till the end of March ...</p> <p>...so I do not see it as being very urgent.</p> <p>I would expect to start looking at it in the New Year.</p> <p>Oliver</p>	<p><b>Gives Orientation</b></p> <p><b>Gives Suggestion</b></p>
	<p><b>Gives Opinion</b></p> <p><b>Gives Opinion</b></p> <p><b>Gives Orientation</b></p> <p><b>Gives Opinion</b></p> <p><b>Gives Suggestion</b></p>

the discussion, the recollection of when expressed.

Steve's contract runs out, followed by the *opinion* based on this that the matter is not urgent. Finally it is *suggested* that the matter is looked at again in the New Year. Note that the language used in both emails does not *ask* for anything.

## **Appendix C – Author’s Related Publications**

### **Guiding Team Selection And The Use Of The Belbin Approach** *International Design Conference 2008*

#### **Abstract**

A collection of successful individuals does not necessarily make a successful team. A comprehensive review of existing methods for forming teams was undertaken, identifying important aspects to consider. In an experiment following fourteen teams, each member’s Belbin roles, academic profiles, skill disciplines and aspects of their tasks were evaluated against team ability. Findings showed the importance of selecting the correct leader and achieving a balance of roles amongst the team. Evidence that different team profiles could be suited to different tasks was also found. Implications for developing a more holistic team selection model are discussed.

WASIAK, J. O., NEWNES, L. B., MILEHAM, A. R., HICKS, B. J. and OUTRAM, G., (2008) Guiding Team Selection And The Use Of The Belbin Approach. IN D. Marjanovic, M. Storga, N. Pavkovic and N. Bojcetic (Eds.) *International Design Conference - Design 2008*. Dubrovnik, Croatia, Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb and The Design Society, Glasgow.



**Characterising The Content Of Email In Engineering Design**  
*Knowledge and Information Management Project Conference 2008*

**Abstract**

Over the last decade there has been a considerable effort by both academia and industry to create improved methods for the organisation and management of the information and knowledge that is required for, used by and communicated between designers and design teams over the product lifecycle.

In general, the information outputs from design projects are formal records, such as reports, computer models and drawings. In addition to these formal outputs, a large amount of potentially valuable information created during the design process is also contained in less formal outputs, such as emails. Because of this potential there is increasing interest in improving the use and management of these informal sources within the context of the overall product lifecycle and the through-life information set. Where emails are considered this includes the creation of:

- improved practices for the construction and formation of emails,
- standards and guidelines for email use, communication and design records,
- methods to support the management, access and reuse of the content of emails,
- approaches to enable the integration of emails with other information sources over the product lifecycle.

However, before such important issues can be addressed there is a fundamental need to understand the information content of emails and their role within the design team and project, a role which until now has been largely overlooked. In order to understand better the role of emails in the engineering context, the research reported in this paper deals with the creation of an approach for assessing the information content of emails and in particular, those associated with design projects. The paper includes a detailed review of the literature and describes the development of an approach for auditing the information content of emails. This unique approach is derived from an extensive evaluation of research in the

domains of computing science, design, social science and management. The approach has been refined iteratively through successive applications to a large email corpus from an industrial design project and is based upon the multi-faceted classification (mark-up) of the content. In particular, the approach considers three perspectives: what the subject of the email is (product, project and organisation), why the email is being sent (problem solving and communication) and how its content is expressed (social interaction and sentiment).

Following a discussion of the development of the approach, its application to an email corpus from a software design project containing more than 700 emails is described. The results reveal how the information content profile changes over the life of the project. In particular it was observed that the amount of product related information was initially low, then grew at an almost exponential rate; whilst the organisational information is higher in the initial project stages but declines as the project progresses. Around 70 percent of the emails analysed contained information relating to the organisation, whereas only 35 percent discussed the product. It was found that a large proportion of emails serve to transfer factual information, without intending a response. Emails were also extensively used as a mechanism for management activities. It was also found that sharing and evaluation of solutions, concerning both the design of the product and more general problems, frequently occurs.

WASIAK, J. O., HICKS, B., NEWNES, L. and DONG, A., 2008. Characterising the Content of Email in Engineering Design. *In* A. Ball and M. J. Darlington (eds.) *KIM Project Conference*, 2nd 3rd April, Univeristy of Reading, UK.

## **Emails In Engineering Design Projects: An Approach For Analysing Their Information Content**

*International Conference on Digital Enterprise Technology 2008*

### **Abstract**

It is increasingly recognised that a large amount of potentially valuable information created during the engineering design process is contained in less formal sources, such as log books and emails. Hence there is a requirement to improve the capture, management and re-use of these sources. However it is first necessary to understand the nature of the information in order to effectively achieve this. The research reported in this paper deals with the creation of an approach for assessing the information content of emails, in particular, those associated with design projects. The development of this approach, grounded in a wide range of existing literature and derived from iterative applications to an industrial dataset is presented. The approach considers three perspectives: what the subject of the email is, why the email is being sent and how its content is expressed. The application of this method to a 700 email corpus from a software design project is described. Key findings from this are presented and discussed, with respect to the character of the content and the project lifecycle.

WASIAK, J. O., HICKS, B., NEWNES, L. and DONG, A., 2008. "Emails In Engineering Design Projects: An Approach For Analysing Their Information Content" *5th International Conference on Digital Enterprise Technology*, 22-24th October, Nantes, France.

**Engineering Work and Email: An Analysis Of The Information  
Content Of Email In A Major Systems Engineering Project**  
*ASME International Design Engineering Technical Conferences &  
Information In Engineering Conference 2009*

**Abstract**

It is said, though not yet assessed, that a large amount of potentially valuable information may reside in engineering email correspondence. If this is the case, then due consideration must be given to the role of email for record keeping, Product Lifecycle Management and knowledge management. In order to examine this hypothesis, a methodology for assessing the information content of emails, and in particular those associated with engineering projects, has been created. The method is based on a textual analysis approach that is derived from cognitive design research and social psychology and further developed through iterative applications to industrial datasets. The paper describes the development of the approach and its classifications of what the subject of the email is, why the email has been sent, and how its content is expressed. The approach is validated using an email corpus from a software design project. The method is then applied to characterize the content of 800 emails from a large systems engineering project. The key findings from this major study are then presented and discussed with respect to the character of the content, including evidence of engineering work, the project lifecycle, and implications for information and knowledge management.

WASIAK, J. O., HICKS, B., DONG, A. and NEWNES, L., 2009. Engineering Work and Email: An Analysis Of The Information Content Of Email In A Major Systems Engineering Project. *ASME IDETC: International Design Engineering Technical Conferences & Information In Engineering Conference*, August 30th - September 2nd San Diego, USA.

## **Managing by e-mail: What e-mail can tell us about engineering project management**

*IEEE Transactions on Engineering Management*

### **Abstract**

E-mails are, rightly or wrongly, a staple of the information and communication technology for managing work and collaborative activities. Yet, the focus of research into the use of e-mail has primarily conceived of e-mail as either a communications tool or a personal information management tool, which ignores the potential value of sharing and reusing privately held information in a wider project or organizational context. This study considers e-mail as part of group information management. The content of e-mails authored by an engineering team associated with a large, complex, long term, systems integration project, typical of the aerospace, marine and defense sectors, was analyzed with respect to why e-mail was sent and its content. The analysis used a methodology developed previously for the investigation of e-mail use in engineering organizations. For the purpose of this study, the results are compared and contrasted with respect to secondary evidence from interviews and project documentation thereby enabling a time-phased analysis of e-mail with respect to the project schedule and the problems experienced. The findings show that e-mails, when analyzed in a group context, reveal predictive signatures of projects that would be useful to project management. In particular, through consideration of the signatures of management, information and problem solving topics, a correlation is revealed between the signatures and the overall project status. Further analysis of the senders and recipients shows that while the e-mail content can identify the roles of senders, these roles are not necessarily consistent with the designated job roles or responsibilities. The findings also raise the problematic issue of balancing the privacy of personal e-mail with the need to conceive of e-mail as part of group information for project management purposes.

WASIAK, J. O., HICKS, B., NEWNES, L., LOFTUS, C., DONG, A. and BURROW, L., n.d. Managing by e-mail: What e-mail can tell us about engineering project management, *IEEE Transactions on Engineering Management*, x (x) x.

**Exploring email use in engineering design projects: a methodology  
based on textual analysis**

*Journal of Research in Engineering Design*

It is widely believed that email is increasingly becoming the medium wherein collaborative engineering work is done; yet, this assumption has not been properly examined. Thus, the extent of engineering information contained in emails and their potential importance within the context of knowledge management is unknown. To address this question, a study was undertaken with a large aerospace propulsion company to investigate the role and characteristics of email communication in engineering design projects. This paper describes the development of a text analysis method for achieving an understanding of email content and hence its use. The technique is based on relevant techniques for analysing communication and design text. The method codes the content of e-mail based on a hierarchical scheme by assigning email to categories and sub-categories that denote what topics the email is about, for which communicative purpose it has been sent, and whether it shows evidence of engineering work. The method is applied to a corpus related to the full life cycle of an engineering design project. Metrics for validation are discussed and applied to a sample case. Exemplar findings are presented to illustrate the type of investigations the method supports. Finally, lessons from the development of the method, including a discussion of iteratively adaptive variants used to arrive at the final outcome, are discussed.

WASIAK, J. O., HICKS, B., NEWNES, L., DONG, A. and BURROW, L.  
2010. Understanding engineering email: the development of a  
taxonomy for identifying and classifying engineering work.  
*Research in Engineering Design*, 21 (1).